Geoscientific
Model Development

Ensemble initialization of the oceanic component of a coupled model through bred vectors at seasonal-to-interannual timescales

J. Baehr and R. Piontek

Institute of Oceanography, Center für Erdsystemforschung und Nachhaltigkeit (CEN), Universität Hamburg, KlimaCampus Hamburg, Germany

Correspondence to: J. Baehr (johanna.baehr@zmaw.de)

Abstract. We evaluate the ensemble spread at seasonal-to-interannual timescales for two perturbation techniques implemented in the ocean component of a coupled model: (1) lagged initial conditions as commonly used for decadal predictions; (2) bred vectors as commonly used for weather and seasonal forecasting. We show that relative to an uninitialized reference simulation the implementation for bred vectors can improve the ensemble spread compared to lagged initialization at timescales from one month up to three years.

As bred vectors have so far mostly been used at short timescales, we initially focus on the implementation of the bred vectors in the ocean component. We introduce a depth-dependent vertical rescaling norm, accounting for the vertical dependence of the variability, and extending the commonly used upper-ocean rescaling norm to the full water column. We further show that it is sufficient for the (sub-surface) ocean to breed temperature and salinity (i.e., scalar quantities), and rely on the governing physics to carry the temperature and salinity perturbations to the flow field.

Using these bred vectors with a rescaling interval of 12 months, we initialize hindcast simulations and compare them to hindcast simulations initialized with lagged initial conditions. We quantify the ensemble spread by analyzing Talagrand diagrams and spread–error ratios. For both temperature and salinity, the lagged initialized ensemble is particularly under-dispersive for the first few months of predictable lead time. The ensemble initialized with bred vectors improves the spread for temperature and salinity for the 0–700 m and 1000–3500 m means, compared to the lagged ensemble at lead times of several months to one year. As the lead time increases to years, the differences between the two ensemble initialization techniques become more difficult to discern. While the results need to be confirmed in an initialized framework, the present analysis represents a first step towards improved ensemble generation at the transition from seasonal to interannual timescales, in particular at lead times up to one year.

1 Introduction

The forecast skill of an ensemble mean is generally more skillful than that of a single deterministic forecast (e.g., Lorenz, 1965; Epstein, 1969; Leith, 1974). An ensemble forecast can provide information pertaining to the likelihood of the occurrence of a particular event, which a single deterministic forecast cannot provide. Hence, ensemble forecasting and the generation of the ensemble has itself been discussed extensively in numerical weather prediction for a variety of different techniques (e.g., Toth and Kalnay, 1997; Hamill et al., 2000; Wang and Bishop, 2003; Keller et al., 2010). For weather forecasting, typically only atmospheric perturbations are required.

At seasonal timescales, it also becomes important to perturb the ocean. Ensemble generation methods are typically implemented with the awareness that ocean perturbations are important, though the effect has mostly been studied for the surface ocean. For example, singular vectors in combination with stochastic optimal wind perturbation (Cheng et al., 2010), perturbing the model coupling physics (Luo et al., 2005) and also the implementation of bred vectors (Cai et al., 2003), have been studied. Beyond two months most techniques showed too little ensemble spread as compared to the forecast error (Vialard et al., 2005).

Predictability at longer than seasonal timescales is thought to reside in the slowly varying components of the climate system. Recently, several studies have shown multi-year predictive skill for example for surface temperature, hurricane frequency or the Atlantic meridional overturning circulation (Smith et al., 2007, 2010; Matei et al., 2012). Additionally, a variety of perfect model studies have analyzed the predictability of the climate system at multi-year timescales (e.g., Pohlmann et al., 2013). However, the issue of ensemble generation has played an ancillary role, and most studies used simple techniques to perturb their hindcast ensemble simulations.

Du et al. (2012) compared the impact of three perturbation strategies, atmosphere only, ocean only, and atmosphere–ocean. They find that for atmospheric variables the spread grows at a similar rate independent of the perturbation method. For ocean-related variables, especially at the subsurface, the spread is sensitive to oceanic perturbations. Du et al. (2012) conclude that any decadal forecast system needs to consider carefully how to perturb the oceanic initial conditions.

Here, we compare two ensemble perturbation methods in the ocean component of a coupled climate model at seasonal-to-interannual multi-year timescales: (1) one-day lagged initialization, a simple ensemble perturbation technique commonly used in multi-year prediction studies (e.g., Müller et al., 2012), and (2) the implementation of bred vectors as developed within the context of numerical weather prediction (Toth and Kalnay, 1993). Breeding is based on the principle that the difference between an unperturbed and a perturbed simulation, i.e., the bred vector, carries information about the fastest growing modes of the model. Bred vectors are regularly rescaled following a pre-defined norm, and subsequently used as initial perturbations for a new round of unperturbed and bred simulations. From a practical standpoint, breeding is an attractive method for ensemble generation as it is relatively simple to implement, yet the resulting perturbations are representative of the model dynamics (Keller et al., 2008).

We implement bred vectors in the ocean component of ECHAM/MPIOM, a coupled climate model that has been used extensively for (perfect model) multi-year prediction experiments (e.g., Pohlmann et al., 2009; Matei et al., 2012). To investigate the impact of oceanic perturbations, we only perturb the ocean component, while the atmospheric component runs freely. We perform perfect model experiments, in which the reference for the breeding is taken to be the unperturbed freely running model. At the expense of being limited to using the model itself as a reference when quantifying the ensemble spread, we avoid the issue of model drift that is commonly associated with initialized simulations (e.g., Pohlmann et al., 2009; Kröger et al., 2012). Such a drift is likely to dominate the variability at short lead times, hence bred vector-generated variability and drift are difficult to separate.

Our breeding implementation focuses on the oceanic component within the coupled model, as well as on the transition from seasonal to multi-year timescales. Previous studies applying breeding in the oceanic part focused on the (near-)surface and sub-seasonal rescaling periods (e.g., Cai et al., 2003; Yang et al., 2006, 2009; Hoffman et al., 2009), though longer breeding cycles have been studied at interannual-to-decadal timescales (Vikhliaev et al., 2007). Here, we therefore initially test the breeding implementation, extending it to include the full water column. After testing the breeding implementation, we compare the ensemble spread of a set of ensemble hindcasts initialized with either bred vector perturbations or lagged initialization. In the ensemble spread analysis, we focus on the critical transition from seasonal to (multi-)year lead times.

2 Model description

We use an updated version of the ECHAM5-MPIOM coupled climate model as used for the IPCC AR4 (Jungclaus et al., 2006). Here, the ECHAM5 atmospheric component (Roeckner et al., 2003) is implemented on a spectral grid at T31 resolution with 19 vertical levels. The MPIOM oceanic component (Marsland et al., 2003) is implemented on an orthogonal curvilinear C-Grid. The northern grid pole is shifted to Greenland, avoiding the singularity at the geographical North Pole. The resulting horizontal resolution is on average about 3°; the vertical resolution is 40 non-equidistant z levels. The atmospheric and oceanic components are coupled with the Ocean-Atmosphere-Sea Ice-Soil (OASIS3) coupler (Valcke, 2006). No flux corrections are applied. The model setup is the same as used in the decadal prediction analysis of Kröger et al. (2012) and Pohlmann et al. (2013), and we use the same model version but in coarser resolution as, e.g., Pohlmann et al. (2009). To avoid model drift, which is commonly associated with hindcast simulations started from an assimilation experiment, we constrain the present analysis to perturbations of the uninitialized freely running model. We use a twentieth century simulation forced with observed greenhouse gas concentrations, and refer to this freely running simulation as our reference simulation. Simulations start from 1 January 1970, and we analyze output averaged to monthly means.

3 Bred vector implementation

3.1 Description

Bred vectors were developed in the 1990s for numerical weather prediction (Toth and Kalnay, 1993), and have been applied extensively within this context (e.g., Toth and Kalnay, 1997). To compute bred vectors, a series of steps is necessary, usually termed the *breeding cycle*. We first

describe the breeding cycle in general terms, and subsequently present our specific implementation:

1. Small random perturbations are added to a reference simulation. Here, we use the difference between the state of the free model at different times at one-month intervals as initial perturbations for the breeding experiments. However, the specific choice of these initial perturbations is not important, as the structure of these initial perturbations is lost after a few breeding cycles are completed (Toth and Kalnay, 1993; Yang et al., 2006).

2. Both the unperturbed reference simulation and the perturbed initial conditions are then integrated forward for a short period of time in which the perturbations grow and propagate.

3. The unperturbed reference simulation is subtracted from the perturbed simulation.

4. The difference between the reference simulation and the perturbed simulation is rescaled using the same procedure as is subsequently used for the breeding itself: the amplitude of the difference is rescaled so that it has the same norm as the variability in the reference simulation.

5. The rescaled perturbations are added to the reference simulation, providing the initial conditions for the new perturbed simulation.

6. Both the unperturbed reference simulation and the perturbed simulation are again integrated forward.

7. The breeding cycle itself consists of steps 3 to 6, repeated iteratively. Here, we allow for a period of 2 yr with monthly normalization until the bred vectors have lost their memory of the initial perturbations, and resemble the fastest growing modes of the model. Subsequently, the breeding experiments consist of 10 ensemble members, each initialized from a different initial perturbation, and are run for 10 yr.

For the implementation of the breeding algorithm the length of the breeding cycle (step 3) and the size of the rescaling norm (step 4 of the breeding cycle) have to be chosen. The exact choice of both parameters depends on the dynamical modes of interest. The length of the breeding cycle should be chosen to isolate the slowly varying dynamical modes (e.g., Pena and Kalnay, 2004; Yang et al., 2006). Typical breeding cycles have a length of several hours in the atmosphere (e.g., Toth and Kalnay, 1993), and month(s) to years in ocean (e.g., Yang et al., 2006; Vikhliaev et al., 2007). For coupled instabilities, the length of the breeding cycle needs to be longer than weather instabilities take to saturate well (Cai et al., 2003). Vikhliaev et al. (2007) use breeding cycles of 6 months and 10 yr to isolate seasonal-to-interannual

and decadal modes of variability in the mid-latitude North Pacific.

The size of the rescaling norm influences the amplitudes of the bred vectors, while it does not influence the structure of the bred vectors (Cai et al., 2003). Usually, the amplitude of the perturbation is calculated for one variable, and the resulting rescaling factor is applied to all variables (Vikhliaev et al., 2007). For coupled instabilities, the norm should be chosen to (also) represent the sub-surface ocean, i.e., measuring the perturbations in the slowly evolving component (e.g., Yang et al., 2006). However, small differences in the choice of the norm do not significantly affect the resulting bred vectors (Vikhliaev et al., 2007).

For the implementation of oceanic bred vectors, the breeding implementation of Yang et al. (2006) has been repeatedly used as a reference setup (e.g., Vikhliaev et al., 2007). Yang et al. (2006) use a one-month breeding cycle and a norm in the equatorial Pacific covering the Niño3.4 region (15° S–15° N, 120° E–90° W). To derive the rescaling factor, the upper ocean temperature (0–200 m) is averaged over the Niño3.4 area. All model variables are then normalized with a constant factor, for example 0.085 °C (Yang et al., 2006). We describe below how we modify this breeding implementation.

3.2 Implementation

Here, we implement bred vectors to perturb the sub-surface ocean for seasonal-to-interannual ensemble prediction experiments. Hence, we implement bred vectors in the oceanic component of the coupled model, while the atmosphere is running freely. Similarly to Yang et al. (2006), we focus on the equatorial Pacific to assess the implementation of the bred vectors. Restricting the implementation of the bred vectors to the oceanic component is likely to prevent the development of bred vectors at shorter timescales. For the seasonal-to-interannual timescale of interest, the ENSO mode is the dominant mode of variability in the tropics, more pronounced in the sub-surface temperature than in the surface temperature (Vikhliaev et al., 2007). For the parameters of the bred vector implementation we make the following choices.

For the length of the breeding cycle, we select 12 months. In the unperturbed reference simulation, anomalies form in the eastern equatorial region of the Pacific basin at the start of summer, and move westward along the equator during the next few months. Hence, the breeding cycle needs to be long enough to allow anomalies to form in the eastern Pacific, and propagate westwards. For short breeding cycles (for example one month), anomalies are allowed to form, but their amplitudes are reduced significantly shortly after they form, and the westward propagation of these anomalies cannot occur (not shown).

For the size of the rescaling norm, we extend the typical mean temperature upper ocean norm to a norm covering

the full water column. As the variability in temperature and salinity significantly decreases at deeper levels, we implement a norm that reflects the decreasing amplitude of variability at depth. This approach is similar to what is used in oceanic state estimates (e.g., Stammer et al., 2002). The bred vectors are normalized individually for each depth level to be ten percent of the root mean square (RMS) profile, ensuring an appropriate amplitude of the perturbation in each layer. The exact choice of the normalization factor is considerably less important than covering the full vertical extent with the norm (not shown).

Instead of applying breeding to all field variables in the model, we limit breeding to temperature and salinity, i.e., to scalar quantities. In test experiments, we find no significant differences between implementing breeding for all field variables (that is, including the velocity field) and restricting breeding to temperature or temperature and salinity (not shown). By restricting breeding to temperature and salinity fields, we rely on the governing physics to carry the temperature and salinity perturbations to the velocity field. We further avoid any potential conflict between perturbations induced in the flow field through breeding the temperature and salinity fields only, and perturbations induced in the flow field through breeding the velocity fields themselves.

3.3 Results

Similarly to Yang et al. (2006), we assess the implementation of the bred vectors by analyzing the spatial structure of the variability of the sea surface temperature (SST) and the upper ocean heat content in the equatorial Pacific. The anomalies of the unperturbed reference experiment show the expected ENSO characteristics in the equatorial Pacific, i.e., an east–west SST gradient (Fig. 1a). The resulting bred vector regression map shows an anomaly in the equatorial Pacific (Fig. 1b), spatially corresponding to the characteristic ENSO structure seen in the unperturbed simulation. The intensity in the central Pacific is slightly weaker than in the unperturbed simulation, and the anomaly does not fully extend to the western boundary, but the essential features of the unperturbed simulation are reproduced. Also, outside of the equatorial Pacific, other structures can be identified in the same locations in both the breeding simulation and the unperturbed simulation (for example in the northern Pacific and in the northern Atlantic).

The regression maps for the temperature averaged over 0–200 m indicate generally similar characteristics (Fig. 1c, d) as the SST (Fig. 1a, b). The unperturbed simulation shows an anomaly from the eastern boundary of the equatorial Pacific into the central Pacific, though weaker than for the SST alone, and negative anomalies occur in the western tropical Pacific (Fig. 1c). The regression map for the bred vectors generally captures these characteristics, although for the bred vectors, the negative anomalies extend further into the central equatorial Pacific than the positive anomalies.
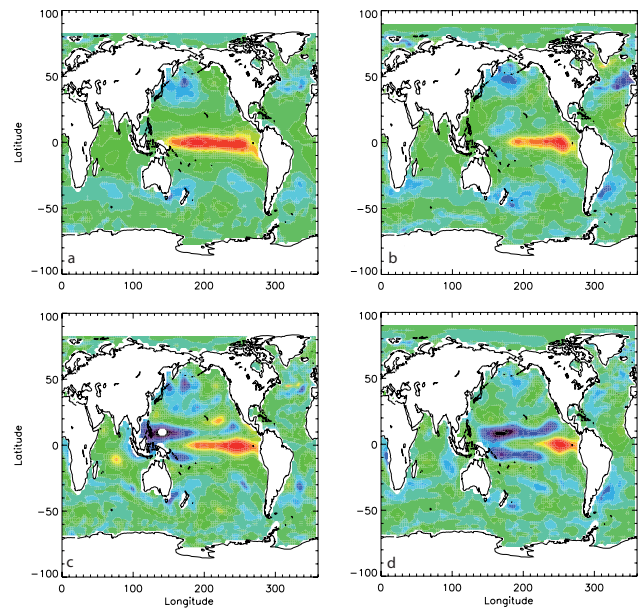

Fig. 1. Regression of the Niño3.4 Index on SST anomalies (top) and 0–200 m temperature (bottom) averaged over the Niño3.4 region for the unperturbed simulation (left; **a** and **c**), and the breeding experiment (right; **b** and **d**). Colors represent the same value in all figures, as normalized values are shown.

As in Yang et al. (2008), we analyze the vertical structure of our bred vector implementation by an empirical orthogonal function (EOF) analysis for the upper ocean temperature across the equatorial Pacific (Fig. 2). The first EOF shows good agreement between the unperturbed and breeding simulations (Fig. 2a, c), accounting for 61 % of the variability in the unperturbed simulation, and 69 % in the breeding. While the first EOF indicates the expected thermocline tilt across the basin, the second EOF shows sub-surface variability in the western Pacific (Fig. 2b, d), accounting for approximately 18 % and 10 % of the variability in the unperturbed run and the bred runs, respectively). The temporal evolution of the principle components for both the first and second EOFs displays the variability of the Niño3.4 Index (not shown). Although the regression patterns for the SST are somewhat weaker in our analysis (Fig. 1b) than in Yang et al. (2006), we find similar EOF structures and explained variability for the upper ocean temperature.

To analyze the robustness of our results, we repeat the analysis both for splitting the analyzed period and also extending it to 20 yr. Neither choice fundamentally changes the regression maps, though the bred vector regression maps show slightly stronger east–west gradients when longer integration periods are analyzed. Similar results are obtained when the number of ensemble members is increased. However, none of these choices fundamentally changes the resulting bred vectors, which resemble the ENSO variability of the reference simulation (not shown).
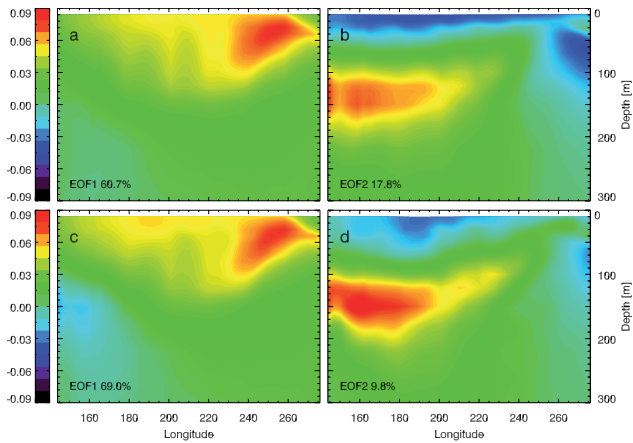
Fig. 2. First and second EOFs of temperature computed from an equatorial slice in the Pacific Ocean for the unperturbed simulation (**a, b**), and for the bred experiment (**c, d**). The numbers on the plots indicate the explained variance of the respective EOF pattern.

In summary, we find a robust representation of the ENSO variability for the following bred vector implementation: a 12-month breeding cycle, a full-depth vertically dependent rescaling norm for temperature, applied to temperature and salinity in the ocean. As a basis for our further analysis, we establish that our bred vector implementation is representative of the model dynamics at seasonal-to-interannual timescales.

4 Ensemble generation

4.1 Description of experiments

We use the bred vectors from the previous section to initialize hindcast experiments. We compare the spread in these hindcast experiments initialized with bred vectors to a set of hindcast experiments initialized with lagged initial conditions. We perform perfect model experiments, in which the reference point is the unperturbed freely running model. Thus we avoid the issue of model drift that is commonly associated with simulations initialized with re-analysis products. The lagged ensembles have a lag interval of one day (as in, e.g., Müller et al., 2012), and only the ocean initial conditions are perturbed. Each bred initialized ensemble member begins with a bred vector perturbation that has previously cycled at least twice. Both bred and lagged hindcasts start in January 1973, with six start dates spaced at one-year intervals. Each hindcast is run out for 10 yr, with ensembles consisting of 10 members. All quantities presented in this section are over six start dates and are computed as a function of lead time. This setup resembles a setup typically used in decadal prediction analyses (e.g., Smith et al., 2007).

As we take the uninitialized run as our reference simulation, we quantify the improvement in the ensemble generation by the evolution of the ensemble spread rather than the predictive skill. We evaluate the evolution of the ensemble spread by two measures: (i) Talagrand diagrams (or rank histograms; Talagrand et al., 1997; Hamill, 2001), and (ii) the spread–error ratio (Palmer et al., 2006; Keller et al., 2008).

(i) For a Talagrand diagram, the ensemble is sorted, and the range of the sorted ensemble verified against a control value, verifying whether the control is within or outside the ensemble distribution. A flat Talagrand diagram indicates that the ensemble has the same probability distribution as the control. A U-shaped Talagrand diagram indicates too little spread in the ensemble, and a Gaussian-shaped rank histogram indicates too much spread. We show Talagrand diagrams for different regions, as distributions vary considerably between different regions. For a Talagrand diagram over a specific region, all grid points and all ensemble members are included individually in the computation of the histogram without any prior spatial averaging. Note that Talagrand diagrams for very small regions might be too noisy to be interpreted for their ensemble spread, while a global average might not representative of the spread in a certain region.

To compare two Talagrand diagrams we compute the deviation of the respective histogram from a flat distribution. We assume variations between two ensembles to be significant if the difference between the two deviations from a flat distribution is larger than 10 percent.

(ii) For the spread–error ratio, we compute the spread as the root mean square difference between the ensemble mean and each ensemble member. The spread–error ratio is then the ratio of this spread and the difference to the reference simulation (here, the uninitialized simulation). An ensemble with the same spread as the reference simulation would result in a spread–error ratio of 1. A spread–error ratio of less than 1 indicates an under-dispersive ensemble, and a ratio greater than 1 indicates an over-dispersive ensemble.

4.2 Results

We first analyze the spread in the global mean upper ocean temperature. While the lagged initialized ensemble is under-dispersive at lead times of one to four months (Fig. 3), the bred initialized ensemble shows a flat distribution in the Talagrand diagram (Fig. 3a). After four months of lead time, the lagged initialized ensemble shows only slightly less spread than the bred initialized ensemble. Similar under-dispersive behavior is found in the spread–error analysis for the global averaged upper ocean temperature (Fig. 3b). After the first months, the spread–error ratio of the lagged ensemble shows mostly values of 0.7 and higher. The spread–error ratio for the bred initialized ensemble shows considerably higher spread–error ratios in the first months as compared to the lagged initialized ensemble. The spread–error ratio for the bred initialized ensemble is on average about 0.05 higher
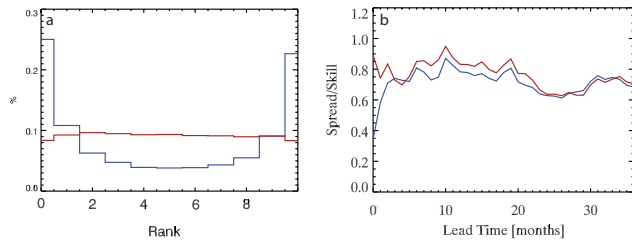
Fig. 3. Globally averaged upper ocean temperature (0–700 m): **(a)** Talagrand diagram for the lagged (blue) and bred (red) initialized ensemble for a lead time of one month. **(b)** Spread–error ratio for the lagged (blue) and bred initialized ensemble (red).

Fig. 4. Talagrand diagrams of the 0–700 m temperature for the lagged (blue) and bred (red) initialized ensembles for lead times of **(a)** one month, **(b)** 2–4 months, **(c)** 6–12 months, **(d)** one year. Histograms are averaged over the grid cell underneath each histogram ($4° \times 8°$ grid from 60° S to 60° N). The shading of the boxes indicates which ensemble is closer to a flat distribution: lagged (blue) or bred (pink) initialized ensemble, or a similar dispersion (green).

than the lagged initialized ensemble out to 24 months lead time.

At a regional scale, differences between the two ensemble initialization techniques are considerably larger than the global mean suggests (Figs. 4 and 5). At one month lead time, the lagged initialized ensemble in the Talagrand diagram is under-dispersive for all considered regions, while the bred initialized ensemble shows a flat histogram in many regions (Fig. 4a). The bred initialized ensemble also shows variability in the shape of the histograms from region to region; the flatness of the global mean is therefore primarily a result of averaging. For lead times up to 12 months, the bred initialized ensemble shows more spread than the lagged initialized ensemble for most regions (Fig. 4b–d).

Similarly, the spread–error ratio indicates considerable regional variation (Fig. 5). At 2–4 months lead time, the lagged initialized ensemble shows a spread–error ratio of less than 1 in most regions (Fig. 5a). The bred initialized ensemble shows a spread–error ratio of about 1 in many regions (Fig. 5b), though several regions with values above/below 1 remain. While Fig. 4 indicates an increased spread in the bred initialized ensemble in many regions, the separate inspection of the error indicates smaller errors in the bred initialized ensemble in the tropics and in dynamically active regions like the western boundary currents.

At a lead time of one year, both ensemble initialization methods show a spread–error ratio closer to 1 for many regions, though the lagged initialized ensemble tends to a spread–error ratio of less than 1, while the bred initialized ensemble tends towards a spread–error ratio of higher than 1 for several regions. At one year lead time, the improvements in the spread–error ratio in the bred initialized ensemble stem almost exclusively from the improvement in the spread, while the errors are of comparable structure and magnitude in both ensembles.

For lead times of two or more years, the differences between the lagged initialized ensemble and the bred initialized ensemble become harder to distinguish (not shown). In the northern and southern Atlantic, the Talagrand diagrams for the bred initialized ensemble tend to be flatter than those for the lagged initialized ensemble, while for most other regions
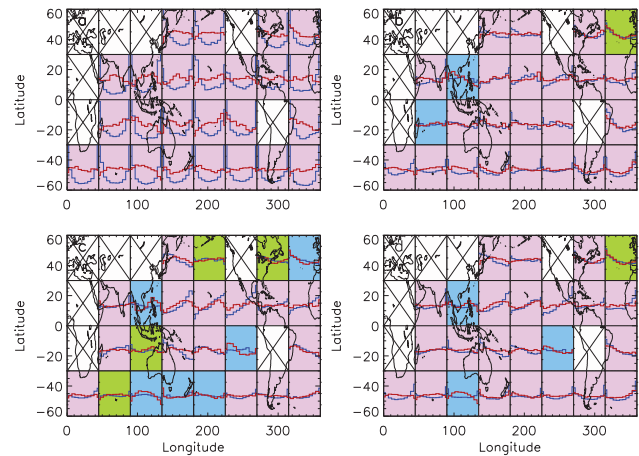
the results for both initialization techniques are comparable. At three years lead time, both initialization strategies show similar spread, which continues for longer lead times (not shown). The spread–error ratio at a lead time of one year is generally closer to 1 for the bred initialized ensemble than for the lagged initialized ensemble, though for both ensemble initialization methods large areas with either over-dispersive or under-dispersive ensembles remain (Fig. 6c, d).

For the deeper ocean temperature, the bred initialized ensemble shows a better spread–skill ratio than the lagged initialized ensemble for nearly all lead times up to three years (Fig. 6a). The Talagrand diagram for the lagged initialized ensemble for one year lead time shows biased ensembles in several regions (Fig. 6b). For several – though not all – of these regions the bred initialized ensemble shows a reduced bias (Fig. 6b). As for the upper ocean temperature, Talagrand histograms indicate comparable spread for the lagged and bred initialized ensembles at longer lead times (not shown).

For upper ocean salinity, both initialization methods yield comparable results for the upper ocean temperature at all analyzed lead times (Fig. 7). The effect of the respective initialization method on the ensemble spread is not only overall similar for temperature and salinity, but is also consistent for most individual regions (Figs. 4a and d, and 7). For one month lead time, the lagged initialized ensemble is considerably less dispersive than the bred initialized ensemble. With longer lead times, the rank histograms for the two initialization techniques converge. However, for lead times up to about two years, the lagged initialized ensemble shows a larger bias and/or less spread than the bred initialized ensemble in most regions.
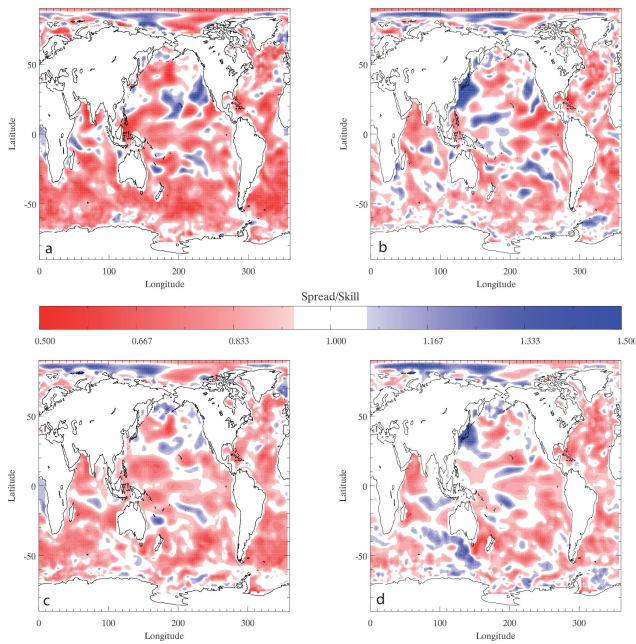
Fig. 5. Spread–error ratios for temperature averaged over 0–700 m depth, and for lead times of 2–4 months **(a, b)**, and one year **(c, d)** for the lagged initialized ensemble **(a, c)**, and the bred initialized ensemble **(b, d)**.

Fig. 6. Spread–skill scores **(a, c, d)** and Talagrand diagram **(b)** for 1000–3500 m temperature: **(a)** globally averaged spread–error ratios for the lagged (blue) and bred initialized ensemble (red), **(b)** Talagrand diagram (as in Fig. 4) for the lagged (blue) and bred (red) initialized ensembles for a lead time of one year; spread–error ratios for one-year lead time for the lagged initialized ensemble **(c)**, and the bred initialized ensemble **(d)**.

Fig. 7. Talagrand diagrams (as in Fig. 4) for 0–700 m salinity for the lagged (blue line) and bred (red line) initialized ensembles for lead times of one month (left), and one year (right).

These results indicate that at seasonal-to-interannual, i.e., one to two years lead time, the bred initialized ensemble shows comparable if not improved spread compared to a lagged initialized ensemble. The lagged initialized ensemble is especially under-dispersive at lead times of a few months. For temperature averaged from 0 to 700 m, the bred initialized ensemble maintains a clear advantage up to lead times of approximately 18 months. For temperature averaged from 1000 to 3500 m, the bred initialized ensemble maintains an advantage out to approximately three years lead time, though the advantage weakens with increasing lead time.

5 Discussion

We implement bred vectors in the ocean component of a coupled model, and perturb the sub-surface ocean by introducing a depth-dependent vertical rescaling norm for the bred vectors. At seasonal-to-annual timescales, the bred initialized ensemble outperforms the lagged initialized ensemble in terms of its spread, and spread–error ratio.

All simulations are conducted in a perfect model framework, and we use an unperturbed control simulation as a reference. We adopt this approach in order to avoid the issue of model drift, which often occurs when initializing with the observed state of the climate system via a re-analysis product. What we present here is a first step, and more experiments in a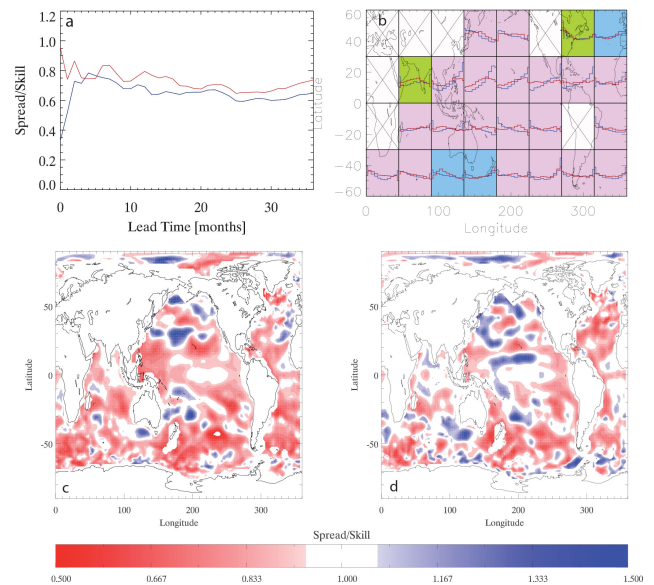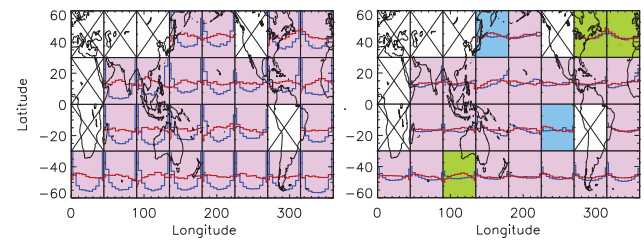n initialized framework, as well as at higher resolution and larger ensemble size, are necessary to transfer the results to a realistic forecast setup.

In the present study, we restrict breeding to the ocean, and let the atmosphere run freely. The assumption is that by perturbing the ocean, we perturb the atmosphere as well, at least at timescales of two weeks and longer. We therefore focus on a one-year breeding cycle, where our simulations yield comparable results to those of Yang et al. (2006), where breeding has been implemented in both the atmosphere and the ocean. Furthermore, we focus in the analysis on oceanic quantities, where we find a more realistic representation of the ensemble spread in the bred initialized ensemble. An improvement in the spread of atmospheric quantities, however, is more likely to be expected when breeding is additionally implemented in the atmosphere. In the present study, we focus

on the different choices of implementing bred vectors in the ocean component, and the sensitivity to the different choices in the bred variables, and the rescaling norm. We maintain that performing the initial analysis in an uninitialized setup where breeding is restricted to the ocean allows us to attribute the effects of the different choices in the bred vector implementation. In a next step, the full depth ocean norm could be combined with breeding in the atmosphere to initialize coupled ocean–atmosphere bred vectors.

6 Conclusions

Based on our analysis of the ocean component within simulations of the ECHAM5/MPIOM coupled climate model, we conclude:

– A breeding implementation restricted to temperature and salinity, relying on the model physics to perturb the velocity field, yields similar results as breeding of all variables that is typically used by Yang et al. (e.g., 2006).

– Using a full watercolumn depth-dependent vertical norm allows us to include the deep ocean in the breeding norm, resulting in a representation of the subsurface instabilities in the bred vectors.

– Analyzing the ensemble spread for individual regions for the two different ensemble initialization techniques yields considerably different results than restricting the analysis to the global mean.

– For seasonal-to-interannual lead time, the bred initialized ensemble shows for most analyzed regions at least similar if not improved spread compared to a lagged initialized ensemble for ocean temperatures averaged over 0–700 m and 1000–3500 m, and also salinity.

Edited by: J. Annan

References

Cai, M., Kalnay, E., and Toth, Z.: Bred Vectors of the Zebiak-Cane Model and their potential application to ENSO Predictions, J. Climate, 16, 40–56, 2003.

Cheng, Y., Tang, Y., Jackson, P., Chen, D., and Deng, Z.: Ensemble construction and verification of the probabilistic ENSO Prediction in the LDEO5 Model, J. Climate, 23, 5476–5497, 2010.

Du, H., Doblas-Reyes, F. J., García-Serrano, J., Guemas, V., Soufflet, Y., and Wouters, B.: Sensitivity of decadal predictions to the initial atmospheric and oceanic perturbations, Clim. Dynam., 7–8, 2013–2023, 2012.

Epstein, E. S.: Stochastic dynamic prediction, Tellus, 21, 739–759, 1969.

Hamill, T. M.: Interpretation of Rank Histograms for verifying ensemble forecasts, Mon. Weather Rev., 129, 550–560, doi:10.1175/1520-0493(2001)129<0550:IORHFV>2.0.CO;2, 2001.

Hamill, T. M., Snyder, C., and Morss, R. E.: A Comparison of Probabilistic Forecasts from Bred, Singular-Vector, and Perturbed Observation Ensembles, Mon. Weather Rev., 128, 1835, doi:10.1175/1520-0493(2000)128<1835:ACOPFF>2.0.CO;2, 2000.

Hoffman, M. J., Kalnay, E., Carton, J. A., and Yang, S.-C.: Use of breeding to detect and explain instabilities in the global ocean, Geophys. Res. Lett., 36, L12608, doi:10.1029/2009GL037729, 2009.

Jungclaus, J. H., Keenlyside, N., Botzet, M., Haak, H., Luo, J.-J., Latif, M., Marotzke, J., Mikolajewicz, U., and Roeckner, E.: Ocean circulation and tropical variability in the coupled model ECHAM5/MPI-OM, J. Climate, 19, 3952–3972, 2006.

Keller, J. D., Kornblueh, L., Hense, A., and Rhodin, A.: Towards a GME ensemble forecasting system: Ensemble initialization using the breeding technique, Meteorol. Z., 17, 707–718, 2008.

Keller, J. D., Hense, A., Kornblueh, L., and Rhodin, A.: On the orthogonalization of bred vectors, Weather Forecast., 25, 1219–1234, 2010.

Kröger, J., Müller, W. A., and von Storch, J.-S.: Impact of different ocean reanalyses on decadal climate prediction, Clim. Dynam., 39, 795–810, 2012.

Leith, C. E.: Theoretical skill of Monte Carlo forecasts, Mon. Weather Rev., 102, 409–418, doi:10.1175/1520-0493(1974)102<0409:TSOMCF>2.0.CO;2, 1974.

Lorenz, E. N.: A study of the predictability of a 28-variable atmospheric model, Tellus, 17, 321, doi:10.1111/j.2153-3490.1965.tb01424.x, 1965.

Luo, J.-J., Masson, S., Behera, S., Shingu, S., and Yamagata, T.: Seasonal climate predictability in a coupled OAGCM using a different approach for ensemble forecasts, J. Climate, 18, 4474–4497, 2005.

Marsland, S., Haak, H., Jungclaus, J., Latif, M., and Röske, F.: The Max-Planck-Institute global ocean/sea ice model with orthogonal curvilinear coordinates, Ocean Model., 5, 91–127, 2003.

Matei, D., Baehr, J., Jungclaus, J. H., Haak, H., Müller, W. A., and Marotzke, J.: Multiyear prediction of monthly mean Atlantic Meridional Overturning Circulation at 26.5° N, Science, 335, 76–79, 2012.

Müller, W. A., Baehr, J., Haak, H., Jungclaus, J. H., Kröger, J., Matei, D., Notz, D., Pohlmann, H., von Storch, J. S., and Marotzke, J.: Forecast skill of multi-year seasonal means

in the decadal prediction system of the Max Planck Institute for Meteorology, Geophys. Res. Lett., 39, L22707, doi:10.1029/2012GL053326, 2012.

Palmer, T., Buizza, R., Hagedorn, R., Lawrence, A., Leutbecher, M., and Smith, L.: Ensemble prediction: a pedagogical perspective, ECMWF Newsletter, 106, 10–17, 2006.

Pena, M. and Kalnay, E.: Separating fast and slow modes in coupled chaotic systems, Nonlinear Proc. Geoph., 11, 319–327, 2004.

Pohlmann, H., Jungclaus, J., Köhl, A., Stammer, D., and Marotzke, J.: Initializing decadal climate predictions with the GECCO oceanic synthesis: effects on the North Atlantic, J. Climate, 22, 3926–3938, 2009.

Pohlmann, H., Smith, D., Balmaseda, M., Keenlyside, N., Masina, S., Matei, D., Müller, W., and Rogel, P.: Predictability of the mid-latitude Atlantic Meridional Overturning Circulation in a multi-model system, Clim. Dynam., 41, 1–11, 2013.

Roeckner, E., Bauml, G., Bonaventura, L., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kirchner, I., Kornblueh, L., Manzini, E., Rhodin, A., Schlese, U., Schulzweida, U., and Tompkins, A.: The atmospheric general circulation model ECHAM5. Part I: Model description, Max-Plank-Report, No. 394, 2003.

Smith, D. M., Cusack, S., Colman, A. W., Folland, C. K., Harris, G. R., and Murphy, J. M.: Improved surface temperature prediction for the coming decade from a global climate model, Science, 317, 796–799, 2007.

Smith, D. M., Eade, R., Dunstone, N. J., Fereday, D., Murphy, J. M., Pohlmann, H., and Scaife, A. A.: Skilful multi-year predictions of Atlantic hurricane frequency, Nat. Geosci., 3, 846–849, 2010.

Stammer, D., Wunsch, C., Giering, R., Eckert, C., Heimbach, P., Marotzke, J., Adcroft, A., Hill, C. N., and Marshall, J.: Global circulation during 1992–1997, estimated from ocean observations and a general circulation model, J. Geophys. Res., 107, 3118–3155, 2002.

Talagrand, O., Vautard, R., and Strauss, B.: Evaluation of probabilistic prediction systems, ECMWF Workshop on Predictability, 1–25, 1997.

Toth, Z. and Kalnay, E.: Ensemble forecasting at NMC: The generation of perturbations, B. Am. Meteorol. Soc., 74, 2317–2330, 1993.

Toth, Z. and Kalnay, E.: Ensemble forecasting at NCEP and the breeding method, Mon. Weather Rev., 125, 3297, doi:10.1175/1520-0493(1997)125<3297:EFANAT>2.0.CO;2, 1997.

Valcke, S.: OASIS3 user guide (prism_2-5), PRISM Support Initiative Report, Technical report, PRISM Support Initiative Report, CERFACS, Toulouse, France, 2006.

Vialard, J., Vitart, F., Balmaseda, M. A., Stockdale, T. N., and Anderson, D. L. T.: An Ensemble Generation Method for Seasonal Forecasting with an Ocean Atmosphere Coupled Model, Mon. Weather Rev., 133, 441–453, doi:10.1175/MWR-2863.1, 2005.

Vikhliaev, Y., Kirtman, B., and Schopf, P.: Decadal north Pacific bred vectors in a coupled GCM, J. Climate, 20, 5744, doi:10.1175/2007JCLI1620.1, 2007.

Wang, X. and Bishop, C. H.: A comparison of breeding and Ensemble Transform Kalman Filter ensemble forecast schemes, J. Atmos. Sci., 60, 1140–1158, 2003.

Yang, S.-C., Cai, M., Kalnay, E., Rienecker, M., Yuan, G., and Toth, Z.: ENSO bred vectors in coupled ocean atmosphere general circulation models, J. Climate, 19, 1422, doi:10.1175/JCLI3696.1, 2006.

Yang, S.-C., Kalnay, E., Cai, M., and Rienecker, M. M.: Bred vectors and tropical Pacific forecast errors in the NASA coupled General Circulation Model, Mon. Weather Rev., 136, 1305, doi:10.1175/2007MWR2118.1, 2008.

Yang, S.-C., Keppenne, C., Rienecker, M., and Kalnay, E.: Application of Coupled Bred Vectors to Seasonal-to-Interannual Forecasting and Ocean Data Assimilation, J. Climate, 22, 2850, doi:10.1175/2008JCLI2427.1, 2009.