https://doi.org/10.5194/egusphere-2026-1340
Preprint. Discussion started: 6 May 2026

Ensemble Generation for Seamless Prediction in the GEOS-S2S Forecast System

Anna Borovikov[1], Andrea Molod[2], Young-Kwon Lim[3,4], Siegfried Schubert[1], and Priyanka Yadav[5,2]

[1]Retired
[2]NASA, Goddard Space Flight Center, Greenbelt, MD 20771
[3]GESTARII, University of Maryland Baltimore County, Baltimore, MD
[4]Environmental Planning Institute, Seoul National University, Seoul, South Korea
[5]ESSIC, University of Maryland, College Park, MD

Correspondence: Andrea Molod (andrea.molod@nasa.gov)

Abstract. Improving the quality of short term climate (subseasonal to seasonal) forecasts depends on improving both the quality of the forecast model and the quality of the initial conditions, with the latter typically consisting of an ensemble of states that are equally likely estimates of the true initial state. In practice, due to our limited knowledge of the true initial errors, an alternative goal is to insure that the initial perturbations project onto the relevant fastest growing modes. With that goal in mind, we present here a relatively simple to implement, yet effective, strategy for generating initial perturbations that are particularly relevant to the short-term climate prediction problem. The strategy, referred to as the Synchronized Multiple Time-lagged (SMT) approach, uses the information about the temporal coherence of nearby analysis states to generate multiple perturbations that are imposed at a specified initial time, with pre-specified amplitudes determined as a fraction of the climatological variance. We show that the perturbations so generated consist of a rich array of physically realistic atmosphere and ocean modes of variability that appear to have some correspondence with the fastest growing modes determined from a singular value decomposition of the model's linear propagator. Furthermore, recognizing the conflicting goals of increasing ensemble size and increasing model complexity, we outline a strategy for reducing, after a specified lead time, an initially large forecast ensemble, which involves performing a stratified sampling of the early larger ensemble in a way that accounts for the emerging directions of error growth.

1 Introduction

Weather and climate forecasts are fundamentally probabilistic in nature. This probabilistic character is based on the fundamental limits of predictability in the earth system, on the need for information about forecast uncertainty, and on the need for information about the probability of occurrence of a forecasted event or state. Many studies over many generations of seasonal prediction systems have been conducted to improve the initial ensemble of states employed in prediction systems (eg., Buizza and Palmer, 1995; Toth and Kalnay, 1997; Sansom et al., 2016; Vitart and Takaya, 2021). The ultimate goal of an ensemble forecast is for the ensemble to be accurate, that is, for the error to be small, and reliable, that is, for the forecast probability correspond to the probability of an event's occurrence. The mean error is largely related to the fidelity of the forecast model

https://doi.org/10.5194/egusphere-2026-1340
Preprint. Discussion started: 6 May 2026

and to the ensemble size, and the reliability to the ensemble perturbations. Perturbation methods are designed to simulate the sources of forecast errors linked to initial condition and model uncertainties, and most methods attempt to ensure that the perturbations project onto the fastest growing modes (eg., Buizza and Palmer, 1995; Yang et al., 2004; Magnusson et al., 2008). This approach is based on the assumption that the fastest growing modes are the ones most likely to degrade the forecasts at longer leads, and has been shown to result in forecasts for which the growth rate in the spread of the ensemble is similar to the forecast error growth rate.

There have been a number of different approaches to generating the initial ensemble. These include (scaled or unscaled) time lagged approaches (e.g., Hoffman and Kalnay, 1983; Dalcher et al., 1988; Kalnay, 2003; Delsole et al., 2017) and a collection of 'burst' approaches that include projections onto both dynamical (e.g., Magnusson et al., 2008) and empirical (e.g., Ham, 2012) estimates of the leading singular vectors, breeding approaches (e.g., Toth and Kalnay, 1997; Yang et al., 2004; Baehr and Piontek, 2014), and model parameter perturbation approaches such as the Stochastic and Perturbed Parameter (SPPT) method of Christensen et al. (2015) or the kinetic energy backscatter approach of Berner et al. (eg., 2015). There are also 'lagged-burst' combination approaches, described in Manrique-Suñén et al. (eg., 2023). Each of these has shown some success in producing what appear to be realistic perturbations to the initial conditions, though each has some drawbacks.

The time lagged approach is relatively easy to implement, but is limited by the length of the lags over which the "perturbations" are sufficiently far from saturation, even when using a scaled averaging approach. The singular vector approaches provides direct estimates of the fastest growing modes, but require either a linearized version of the GCM (dynamical) or long history of forecasts to estimate the (empirical) linear operator. The breeding approach also aims to identify the "errors of the day" using the full GCM to grow errors, but requires substantial additional computing for the breeding.

Here we describe an alternative approach to generating initial perturbations, called the 'Synchronized Multiple Time' (SMT) approach, that is relatively easy to implement yet at the same time appears to be very effective in generating perturbations that reflect a broad range of weather and climate modes of variability, and yet is not limited by the length of the time lags. As such SMT can provide realistic estimates of how the errors of the day can impact the weather and climate modes at longer leads.

Recent studies have demonstrated the role of ensemble size in forecast accuracy (Scaife and Smith, 2018), particularly at shorter lead times when the predictability associated with tropical-extratropical teleconnections is relevant. Given the conflicting goals of large ensembles and increased model complexity, we also outline here a strategy for reducing, after a specified forecast lead time, an initially large forecast ensemble. This ensemble size reduction involves performing a stratified sampling of the early larger ensemble in a way that accounts for the emerging directions of error growth.

In addition to a description of SMT and the clustering strategy, we also provide here an evaluation of the elements of both, and of the impact on the relationship between the spread and error in the GEOS-S2S-3 suite of retrospective forecasts in which SMT was implemented.

This paper is a follow-on of the work described in the Schubert et al. (2019) technical memorandum, with additional new results based on the new Goddard Earth Observing System Subseasonal to Seasonal forecast system, version 3 (GEOS S2S-3), in which SMT has been implemented. Section 2 describes the GMAO GEOS S2S forecast system. Our approach to generating the ensemble is described in Section 3.1, and our approach to subsampling is presented in Section 3.2. In section 4.1 we present

https://doi.org/10.5194/egusphere-2026-1340
Preprint. Discussion started: 6 May 2026

an example of the impact of stratified sampling based on a single ensemble of side by side retrospective forecasts with GEOS-S2S-2. Our initial assessment of the perturbation error growth in the current suite of GEOS S2S-3 retrospective forecasts is presented in Section 4.2. A summary and set of conclusions is presented in Section 5.

2 The GEOS S2S Forecast System

Previous versions of the GEOS-S2S system are described in Borovikov et al. (2017) (GEOS-S2S-1, used in near-real time from 2012 to 2018) and Molod et al. (2020) (GEOS-S2S-2, used in near-real time from 2018 to 2025). GEOS-S2S-3 is being used for near-real time assimilation and forecasts, and to generate a 1999-present coupled reanalysis. The GEOS-S2S system includes the coupled Atmosphere-Ocean General Circulation Model (AOGCM), the coupled Data Assimilation System (AODAS), and the forecast initialization and ensemble perturbation approaches. We present here an overview of GEOS S2S-3, highlighting similarities and difference from GEOS-S2S-2, our baseline system which was used to conduct many experiments during the development of the new ensemble strategy.

The main components of the GEOS AOGCM are the GEOS AGCM (Molod et al., 2015; Rienecker et al., 2008), the catchment land surface model (Koster et al., 2000), the GOCART aerosol model (Chin et al., 2002; Colarco et al., 2010), the MOM5 ocean general circulation model (Griffies et al., 2005; Adler et al., 2003; Griffies, 2012), and the Community Ice CodE-4 sea ice model (Hunke and Lipscomb, 2008). In GEOS-S2S-3, the atmosphere and ocean are coupled using the Atmosphere-Ocean Interface Layer (AOIL) of Akella and Suarez (2018), and the meltwater runoff from ice sheets was improved to create discharge in the proper location. In addition, the GEOS-S2S-3 AOGCM underwent a major retuning exercise using the Green's functions methodology described in Strobach et al. (2020). All components are coupled together using the Earth System Modeling Framework (Hill et al., 2004) and the Modeling Analysis and Prediction Layer interface layer (Suarez et al., 2007). The atmospheric data assimilation component is the pre-existing "GEOS for Instrument Teams" near-real-time assimilation (GEOS_IT), and the ocean data assimilation follows the Local Ensemble Transform Kalman Filter (LETKF) of Penny et al. (2013). The GEOS-S2S-3 coupled assimilation (called 'GiOCEAN') and forecasts include an AGCM component run on a cubed sphere grid with 0.5° horizontal resolution and 72 hybrid sigma/pressure levels, and the MOM5 ocean components run on a tripolar grid at 0.25° horizontal resolution with 50 layers down to 4500 m depth. The LETKF ocean analysis is run on a lat/lon grid of 0.25° resolution.

The data assimilation coupling methodology (See Fig 1 of Molod et al., 2020) in GEOS-S2S-3 mimics that of GEOS-S2S-2 with a few key upgrades, and includes both an ocean predictor segment and a corrector segment. During both segments, the atmospheric state is "replayed" using a "regular replay" (Orbe et al., 2017; Takacs et al., 2018) (GEOS-S2S-2 employed an 'intermittent replay') to GMAO's GEOS_IT atmospheric near-real time assimilation (GEOS-S2S-2 replayed to the Forward Processing for Instrument Teams, FPIT, assimilation). "Replay" is a form of nudging, wherein the atmospheric state is constrained to approximate the atmospheric analysis. The replay is performed to constrain atmospheric state variables (winds, temperqture, humidity and mass) and aerosol fields (Buchard et al., 2017) (GEOS-S2S-2 did not replay to aerosol fields). During both segments of the AODAS the land model is forced using observed CMAP precipitation (Reichle and Liu, 2014; Xie and

https://doi.org/10.5194/egusphere-2026-1340
Preprint. Discussion started: 6 May 2026

Arkin, 1997) rather than the model's predicted precipitation. After the 5-day predictor segment, the ocean analysis increments are computed using the LETKF and assimilating a full suite of *in situ* and remote ocean observations, and the coupled AODAS returns to the beginning of the 5-day segment to perform the corrector segment using the incremental analysis update (IAU) method of Bloom et al. (1996). For all the GEOS-S2S systems to date, the forecasts are initialized on a fixed set of calendar dates, beginning on 1 January and each 5 days hence. This choice of start dates is to ensure that the ocean observations in the 5-day data window are absorbed by the AODAS and that the earth system has achieved dynamical balance at the start of the forecast segment.

In the default setup for GEOS-S2S-2 the retrospective and near-real-time seasonal forecasts are initialized on the last four start dates of the month. For the retrospective suite, no additional ensemble members are generated, while for the near-real-time forecasts an additional six ensemble members are generated by perturbing the state of the last start date in each month, thus producing a total of 10 ensemble members with this "lagged-burst" strategy. The method to perturb initial conditions is based on the difference between two analysis states 5 days apart. The perturbations are rescaled and the magnitude of the norm reduced to approximately 10% of the natural variability of SST over the norm region, 120–90W, and 10S to 10N. The variables perturbed on the ocean model grid are temperature, salinity, and ocean velocities, and the variables perturbed on the atmospheric grid are wind components, potential temperature, specific humidity and mass. The motivation for improvement of this ensemble perturbation strategy and the development of the new strategy detailed here is described in Schubert et al. (Section 2.4 of 2019).

The retrospective (and near-real time) GEOS-S2S-3 forecasts are initialized from "GiOCEAN" (GiOcean-NRT) employing the new perturbation/ensemble generation technique that is summarized in Section 4.2.

3 The Ensemble Generation and Sub-sampling Techniques

As mentioned in Section 1, most ensemble generation strategies are a designed to produce perturbations that project onto the fastest growing disturbances. For the short-term climate prediction problem, it is important to project directly onto modes that contain the predictability on subseasonal to seasonal time scales such as the Madden-Julian Oscillation (MJO), El Niño-Southern Oscillation (ENSO) and teleconnections. Our approach (described below) is a simple 'lagged-burst' technique that we show closely approximates more computationally intensive methods, and is designed to inject uncertainty in all the modes of interest. This allows for the fact that different modes may be active at different times and as such, addresses the "errors of the day" without explicitly isolating those errors. We refer to this approach (described next) as the Synchronized Multiple Time-lagged (SMT) technique. In this section we describe both the SMT technique (Section 3.1) and our approach to subsampling the ensemble (Section 3.2). Additional details can be found in Schubert et al. (2019).

3.1 The Synchronized Multiple Time-lagged (SMT) Technique

The SMT technique computes perturbations as scaled differences between two nearby analysis states for various separations in time. The differences are scaled to ensure that the perturbations are of the order of some specified fraction of the climatological

https://doi.org/10.5194/egusphere-2026-1340
Preprint. Discussion started: 6 May 2026

variance of the perturbed quantity. The SMT technique, similar to the technique used in GEOS-S2S-2, initializes forecasts every 5 calendar days (the 'lagged' component), and calculates perturbations (the 'burst' component) as scaled analysis differences. In SMT, however, the perturbations are computed from a random selection of separation days ranging from 1 to 10 days, and from a random selection of dates over the last month from which to compute the differences. The scaling developed for SMT is described here and in Schubert et al. (2019). The scaling is a function of the temporal autocorrelations of the differences, allowing control over the amplitudes of the various perturbations based on the 'errors of the day'. The scaling is also designed to produce perturbations that are small relative to the temporal variations of the quantity in question, thus retaining the dynamical balance of the unperturbed state. Understanding the spatial structure of the perturbations and the relationship to the fastest growing modes requires an examination of the spatial covariances of the differences and how they vary as a function of the time separation.

3.1.1 Autocorrelations and Scaling

The derivation of the expression for the scaling of the analysis differences at each separation time τ is provided in Schubert et al. (2019), and is designed to ensure that the perturbations have the same magnitude relative to climatology, independent of the separation τ. The scaling factor is written here as:

$$\alpha_\rho(\tau) = \frac{\epsilon}{(2 * (1 - \rho(\tau)))^{1/2}} \tag{1}$$

where $\rho(\tau) = \mathrm{Corr}(\mathrm{X(t+\tau)}, \mathrm{X(t)})$ is the autocorrelation based on the daily data and ϵ is whatever fraction of the climatological standard deviation one wants the magnitude of the perturbations to be (say 0.1). We note that at long separations ($\rho(\tau) \rightarrow$ zero) the scaling factor reduces to $\frac{\epsilon}{2^{1/2}}$.

Figure 1 shows some examples of the scaling factor (with ϵ = 0.1) for atmospheric potential temperature, specific humidity, and the zonal and meridional components of the wind. A key result is that the values generally decay to the limit of $\frac{\epsilon}{2^{1/2}} = 0.07$ after about one week of separation, consistent with the typical decorrelation time scale of large-scale atmospheric teleconnections (e.g., Feldstein, 2000). The other key result is that the values for different fields at different atmospheric levels are overall rather similar to each other, though there is clearly a longer memory closer to the surface for the temperature and moisture fields. We can take advantage of this result and produce an average set of scaling factors that vary only with season and separation as shown in Figure 2.

Figure 3 shows the results for the ocean temperature and salinity at various depths. While the overall time scales for the ocean are clearly longer than they are for in the atmosphere, our interest here is again on short term error growth and so we again focus on differences in the ocean state just a few days apart. The results do show somewhat larger values of the scale factor compared with the atmosphere especially for the 1-day differences, consistent with the longer ocean time scales. While there is considerable scatter in the scale factors at the 1-day separation as a function of depth, it is likely sufficient in practice to utilize a single averaged set of scale factors that depend only on the separation τ.

https://doi.org/10.5194/egusphere-2026-1340
Preprint. Discussion started: 6 May 2026

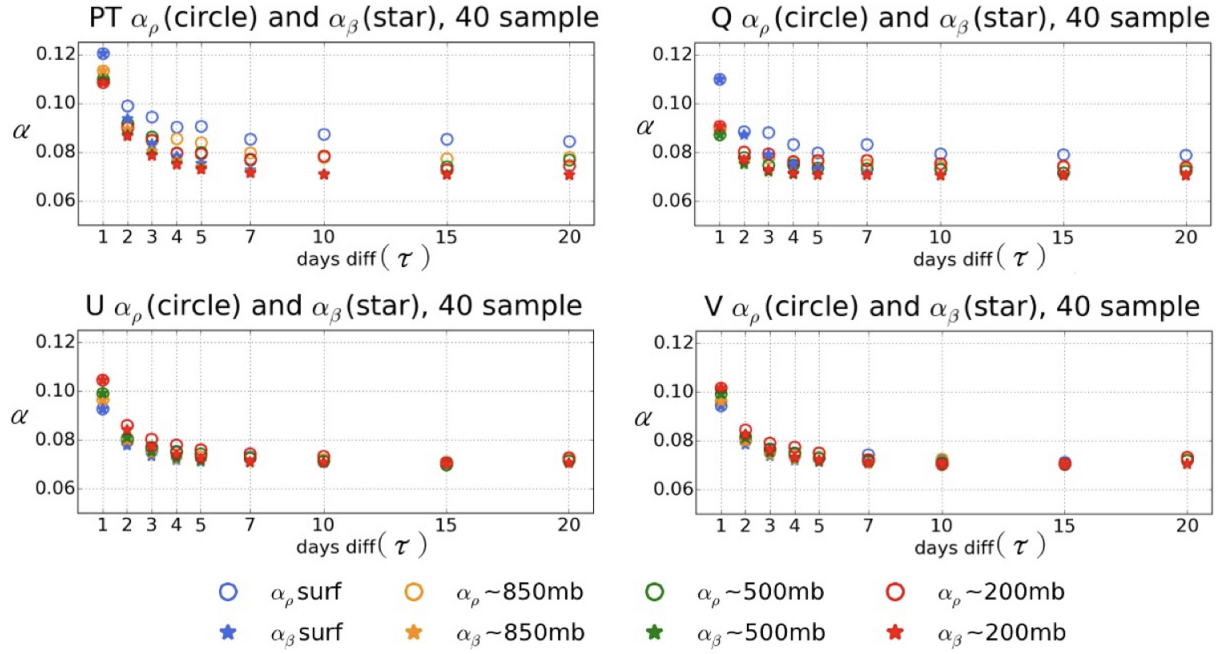

Figure 1. The perturbation scaling factor $\alpha_\rho(\tau)$ for various atmospheric quantities and levels as a function of separation τ in days. Top left: potential temperature, top right: specific humidity, bottom left: u-wind and bottom right: v-wind. Also shown is $\alpha_\beta(\tau)$ which assumes the autocorrelations follow a simple first order autoregressive process. Results are based on a sample of 40 differences taken from coupled data assimilation system (AODAS) restarts during the period October-November-December 2017.

3.1.2 Spatial covariances

Here we examine the spatial structure of the perturbations computed as a difference between two analysis states, with a particular interest in how the structures vary as a function of the separation. This analysis is demonstrated in Schubert et al. (2019), but repeated here to emphasize the connection between these perturbation structures and the fastest growing error modes that they are designed to capture. We begin by assuming that the perturbations (which can be considered as a tendency in time or separation) are approximately governed by:

$$\Delta \boldsymbol{X}_{\boldsymbol{\tau}}(t) = \boldsymbol{X}(t+\tau) - \boldsymbol{X}(t) \approx \mathbf{A}_\tau \boldsymbol{X}(t) \tag{2}$$

where $\mathbf{A}_\tau$ is an n x n matrix and $\boldsymbol{X}$ is an n x 1 vector representing the daily anomalies (over say the n grid points) of the climate system. Here the anomalies are defined as deviations from a mean (< >), where the angle brackets denote an average over the history of perturbations being considered (either a long history over many years of perturbations, or a recent history of perturbations just prior to the start of the particular forecast in question). Note that the linear propagator, $\mathbf{A}_\tau$, depends on

https://doi.org/10.5194/egusphere-2026-1340
Preprint. Discussion started: 6 May 2026

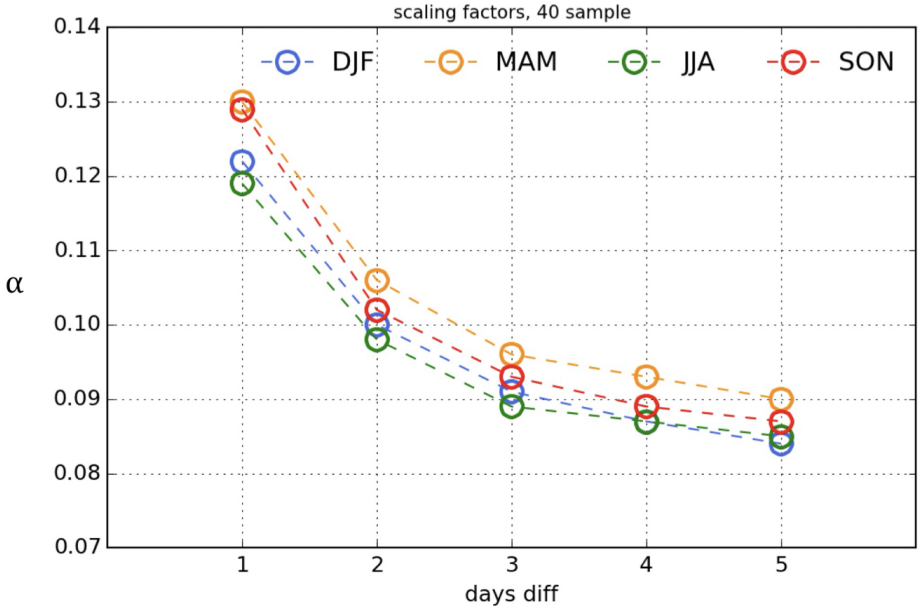

Figure 2. The overall average atmospheric perturbation scaling factor $\alpha_\rho(\tau)$ for each season as a function of separation τ in days. Results are based on a sample of 40 differences taken from coupled data assimilation system (AODAS) restarts.

the lead time τ. Since we are interested in the spatial structure of the perturbations $\Delta \boldsymbol{X_\tau}(t)$ as a function of τ we examine the covariance matrix:

$$\mathbf{D}_\tau = \langle \Delta \boldsymbol{X_\tau}(t)\, \Delta \boldsymbol{X_\tau}(t)^T \rangle \tag{3}$$

where the superscript T denotes a matrix transpose. Substituting for $\Delta \boldsymbol{X_\tau}(t)$ from equation 2 we obtain:

$$\mathbf{D}_\tau = \mathbf{A}_\tau \langle \boldsymbol{X}(t)\, \boldsymbol{X}(t)^T \rangle \mathbf{A}_\tau{}^T = \mathbf{A}_\tau \Sigma \mathbf{A}_\tau{}^T \tag{4}$$

where $\Sigma = < \boldsymbol{X}(t)\, \boldsymbol{X}(t)^T >$ is the covariance matrix of the daily data.

This shows that our perturbations (to the extent that they reflect the leading EOFs of $\mathbf{D}_\tau$) are closely related to the optimal perturbations that would be obtained from a singular value decomposition (SVD) of the operator $\mathbf{A}_\tau$. Here the singular vectors of $\mathbf{A}_\tau$ are the eigenvectors of the matrix $\mathbf{A}_\tau\, \mathbf{A}_\tau{}^T$ (Strang, 2009).

If we assume that $\Sigma = \mathbf{I}$ (ie, our initial conditions $\boldsymbol{X}(t)$ are uncorrelated white noise in space) we would be sampling from a covariance matrix that has eigenvectors identical to the left singular vectors of $\mathbf{A}_\tau$. Schubert et al. (2019) describes the more general case when $\Sigma \neq \mathbf{I}$. In addition, since the eigenvalues of $\mathbf{D}_\tau$ are just the square of the singular values, we would presumably be sampling preferentially those perturbations with the largest growth rates. As such, we would expect that

https://doi.org/10.5194/egusphere-2026-1340
Preprint. Discussion started: 6 May 2026

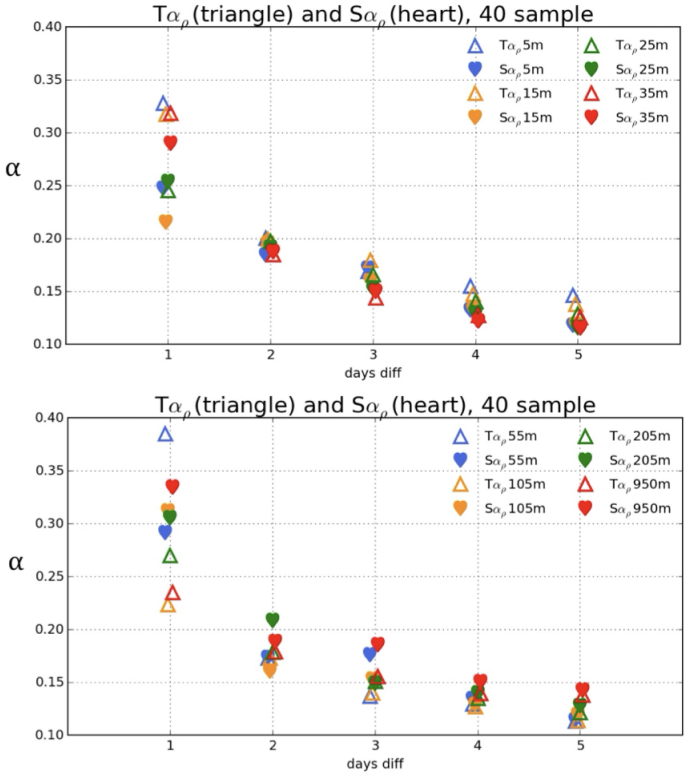

Figure 3. The perturbation scaling factor $\alpha_\rho(\tau)$ for ocean temperature (T) and salinity (S) as a function of separation τ in days at various depths. Top: for shallow depths (5, 15, 25, 35 meters). Bottom: for deeper depths (55, 105, 205, 950 meters). Results are based on a sample of 40 differences taken from coupled data assimilation system (AODAS) restarts during the period OND 2017.

the leading EOFs of $\mathbf{D}_\tau$ resemble the fastest growing disturbances as determined from a singular value decomposition of $\mathbf{A}_\tau$ (with some dependence on the prevailing covariance structures as determined by Σ), and those leading EOFs would likely change as a function of the separation since $\mathbf{A}_\tau$ is itself a function of τ. This indicates that there is something to be gained by creating perturbations based on several different values of τ, in that it allows us to perturb different modes of variability that contribute to forecast uncertainty on different time scales and in different regions of the globe.

As an example, we show in Figure 4 the leading EOFs of the mid-tropospheric potential temperature at 1-day and 5-day differences for September – November. The differences in the spatial structure of the leading EOFs at the 1-day differences (left panel) have a synoptic-scale structure typical of NH middle latitude weather systems, while the 5-day differences produce larger-scale teleconnection patterns reminiscent of, for example, the Pacific/North American pattern (PNA, Wallace and Gutzler, 1981). The leading EOFs at τ 1 through 5 days represent anywhere between about 8% and 13% of the total global variance of the difference fields (not shown).

https://doi.org/10.5194/egusphere-2026-1340
Preprint. Discussion started: 6 May 2026

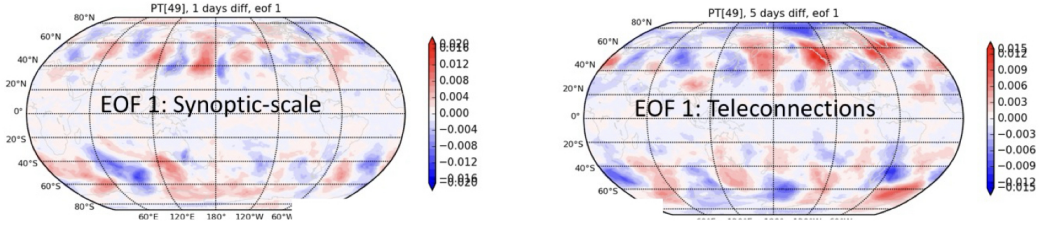

Figure 4. Typical structures of atmospheric perturbations in middle latitudes for potential temperature at model level 49 (approximately 450mb). The patterns are the leading EOFs computed from 1-day (left) and 5-day (right) differences of AODAS restarts during the period September-November (SON) 2017.

While the above EOFs naturally isolated the leading structures in the middle latitudes where the day-to-day variability in temperature is largest, we can also look at what the leading structures are in the tropics by simply confining the domain of the EOF calculation to the tropical region. Figure 5 shows the leading EOFs of the tropical zonal wind at 850mb and 200mb (computed as combined EOFs), for τ = 5 and 10 days during boreal winter. At both 5- and 10- days separation the structures are larger in scale than in the extratropics, with the 10-day separation apparently isolating MJO-like variability. It is noteworthy that very similar results are found for boreal summer (Schubert et al., 2019).

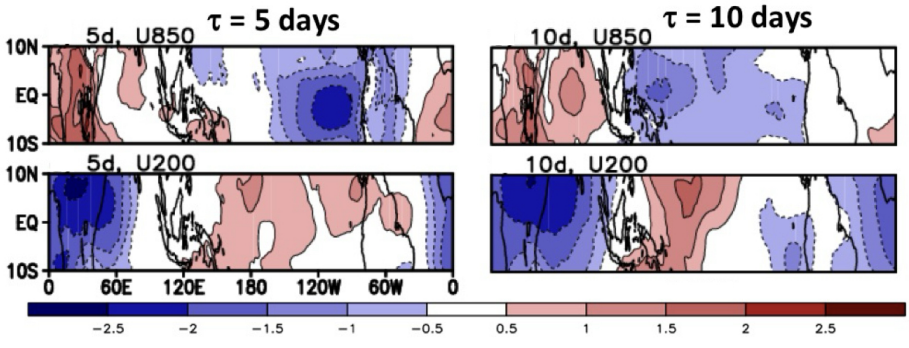

Figure 5. Typical structure of atmospheric perturbations in the tropics for zonal wind at 850mb and 200mb during December, January and February (DJF). The patterns are the leading EOFs computed from 5 day (left) and 10 day (right) differences. Results are based on MERRA-2 for the years 1999-2016.

We next turn to the ocean. As an example, we show in Fig. 6 the leading EOFs of the longitude-depth (x-z) cross-section of the Pacific equatorial temperature perturbations extending to a depth of 300m during September-November for 1-day and 10-day separations. The results show modes in which the variability is to a large extent tied to variations in the thermocline. At one day separation (upper left panel) the leading EOF has values generally of the same sign throughout the mixed layer and below, with particularly large contributions at and just west of the dateline below 100 meters. The correlations with SST

https://doi.org/10.5194/egusphere-2026-1340
Preprint. Discussion started: 6 May 2026

(lower left panel) indicate that this mode is associated with SST variations that are spatially coherent throughout much of the tropical Pacific east of 150E. At 10-days separation the leading EOF again shows relatively large amplitude below 100 meters (though now somewhat further west, near 150E), but there is also clear evidence of a vertically coherent wave-type variability extending throughout the mixed layer in the eastern Pacific. In fact, the correlations with SST in the tropical Pacific (bottom right panel of Fig 6) exhibit a wave structure consistent with that of tropical instability waves (e.g., Shinoda et al., 2009). It is noteworthy that the heat transport associated with tropical instability waves appears to play a critical role in generating the asymmetries between El Niño and La Niña events (e.g., Imada and Kimoto, 2012), hence their likely relevance to the seasonal prediction problem.

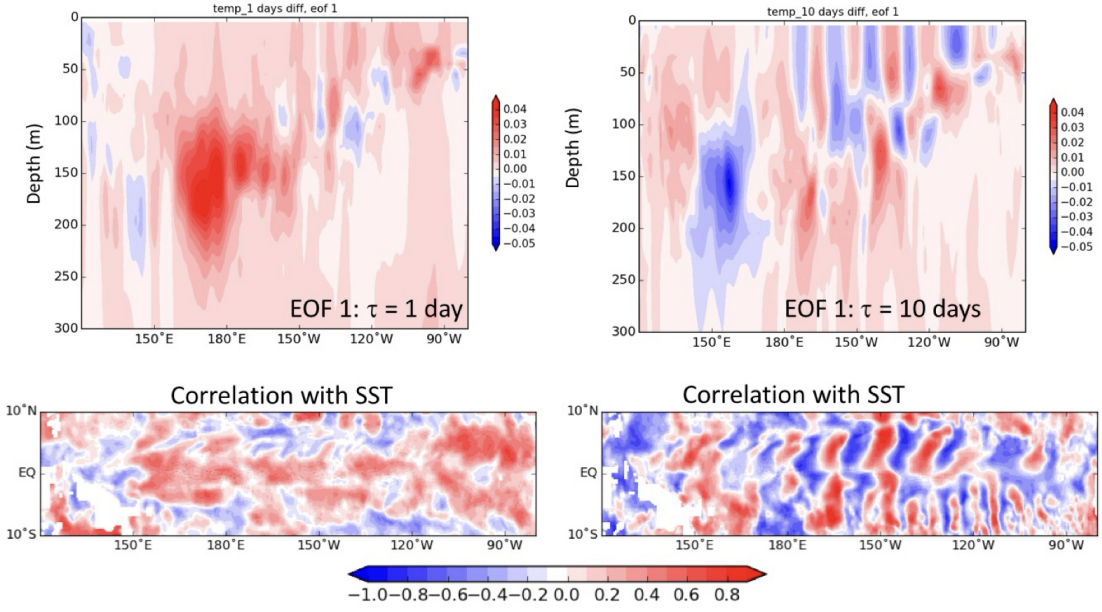

Figure 6. Typical structure of ocean temperature perturbations at τ equals 1-day (top left) and 10-day (top right) separations. The structures are the leading EOFs of the Pacific equatorial x-z cross section of temperature averaged between 2S-2N, and extending down to a depth of 300 meters. The bottom panels show the correlations of tropical Pacific SST with the leading EOF for 1-day separation (left) and 10-day separation (right). Results are based on 100 randomly chosen pairs of AODAS restarts taken from SON of 2017.

The above results suggest that an SMT approach in which the perturbations are based on separations of at least 1, 3, 5 and 10 days would introduce a reasonable blend of different physically realizable error structures. In the atmosphere these include middle latitude baroclinic waves (τ=1 day) and teleconnections (τ=5 days), and MJO-like variability (τ=10 days) in the tropics. In the tropical ocean, the analogous perturbations appear to be tied to variations in the thermocline, with the longer separations

https://doi.org/10.5194/egusphere-2026-1340
Preprint. Discussion started: 6 May 2026

(e.g., τ=10 days) showing evidence of tropical instability waves. These atmospheric and oceanic structures are likely to have a substantial impact on the early (1-2 month) growth of forecast errors in both the ocean and atmosphere. Another important feature of the SMT technique is the inclusion in a single ensemble perturbation system of modes relevant for predictability at both subseasonal and seasonal forecast lead times. This allows GEOS-S2S-3 to be a truly seamless prediction system, unlike GEOS-S2S-2 and also unlike other subseasonal to seasonal prediction systems.

3.2 Sub-sampling the Ensemble

Obtaining improved estimates of the mean as well as the various other probabilistic measures of forecast quality involving higher order moments require that we produce ever larger ensembles. This, of course, has to be weighed against the desire to increase model resolution and complexity in what is, for practical reasons, always a limited computational resource environment.

With that in mind, we develop here a strategy of running with a relatively large number of ensemble members for the shortest forecast leads, during which ensemble size matters (eg., Scaife and Smith, 2018) and beyond which predictability of certain modes decays, and then selecting a smaller subset to continue out to the longest leads where ensemble size is less relevant. The general approach is illustrated schematically in Figure 7. The question addressed here is how to optimally select that subset, taking advantage of well-known results on stratified sampling (Cochran, 1963).

The basic idea is that we take advantage of the information about the early error growth that can be obtained from the relatively large initial ensemble, in a way that ensures the capture, through proper subsampling, of the leading directions (in phase space) of error growth (Schubert et al., 1992). This can be especially important when the ensemble is characterized by more than one dominant direction of error growth.

The approach we use assumes that for a quantity y, we have a large ensemble of size N (the population) that we wish to subsample with n ensemble members, where N$\gg$n. The population of size N is divided into L disjoint strata, where $\mathrm{n_h}$ ($\mathrm{N_h}$) are the number of members of the sample (population) in stratum h. Although Neyman allocation (Cochran, 1963) is optimal, the relatively small sample size suggests that the simpler proportional sampling, wherein each stratum is sampled in proportion to its representation in the population, should be used here, where:

$$\frac{n_h}{n} = \frac{N_h}{N} \tag{5}$$

The success of the sampling strategies requires that we are able to divide the population (N) into L strata. One reasonable approach to doing this is to divide the population in a way that minimizes the average intra-stratum variance (see e.g., Schubert et al., 1992). Here we use the KMEANS clustering algorithm (Spaeth, 1980). This algorithm requires specifying the number of clusters (L) along with some initial guess of the cluster distribution. The details of our KMEANS clustering choices as well as a description of the evaluation of alternatives is articulated below.

https://doi.org/10.5194/egusphere-2026-1340
Preprint. Discussion started: 6 May 2026

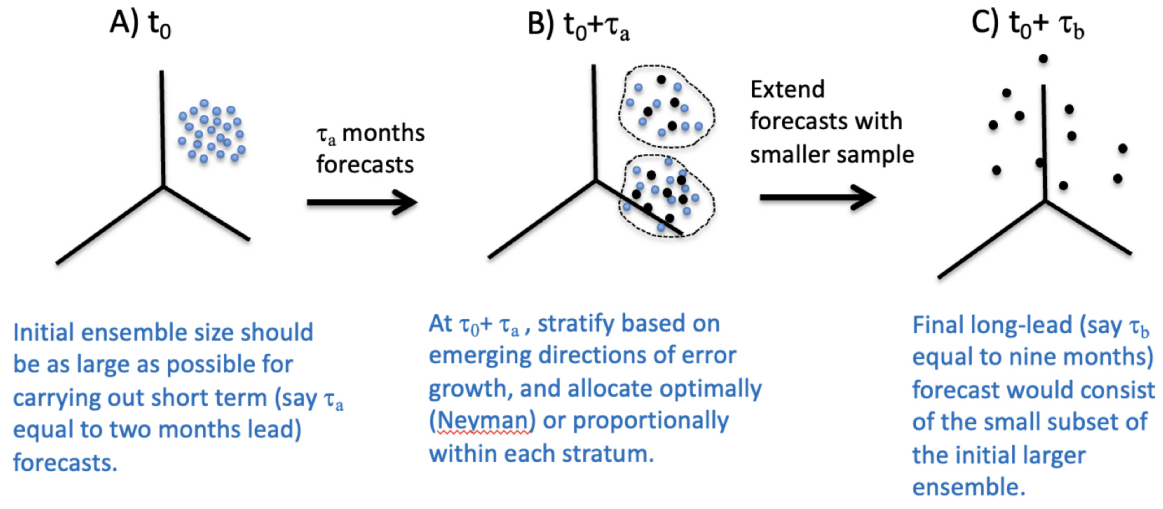

Figure 7. Schematic of stratified sampling of ensemble forecasts.

4 Results of SMT, Subsampling and Growth of Ensemble Spread

The two major aspects of the new ensemble strategy are the effectiveness of the subsampling strategy to ensure that the ensemble spread is retained by the smaller sample, and the ability of the ensemble to properly capture the fastest growing modes, as represented by the relationship between the ensemble spread and ensemble mean error. Both of these aspects are evaluated in this section with a relatively large ensemble of a single month's forecast, using GEOS-S2S-2. Computational constraints determined the use of a single month's ensemble for the side by side forecasts required for this evaluation.

4.1 Impact of Subsampling with GEOS S2S-2

Here we provide an assessment of the subsampling technique by using results of a relatively large (110 members) ensemble of 5 months-long forecasts initialized in November 2017 with the GEOS S2S-2 forecasting system (Section 2, Molod et al., 2020). The full 110-member ensemble consists of 30 that are based on lagged initial conditions (runs started each day in Nov 1-30) and 80 that are initialized on Nov 27 using the burst approach; of the latter, 40 have atmospheric perturbations and another 40 have oceanic perturbations (see Table 1). Data are available at Borovikov (2018).

https://doi.org/10.5194/egusphere-2026-1340
Preprint. Discussion started: 6 May 2026

No. Members	Perturbed Variables	Notes	Initialization Date
40	potential temperature (PT) specific humidity (Q) zonal wind (U) meridional wind (V)	Atmos. perturbations τ = 1, 2, 3, 4, 5 days (8 of each) ϵ = 0.10	11/27/2017
40	temp, salt u, v t_surf, s_surf, u_surf, v_surf sea_lev, frazil	Ocean perturbations τ = 1, 2, 3, 4, 5 days ϵ = 0.10	11/27/2017
30	The entire state	Implicit perturbations via time lag initialized each day	Nov. 1-30, 2017

Table 1. Experiments carried out with S2S-2 to examine the impact of subsampling.

Given the importance of ENSO at the longer forecast leads, we focus on the Niño3.4 SST index. We treat the N=110 ensemble members as our population and wish to choose an optimal subsample in the sense that it minimizes the variance of the sample mean. In this example we use an n=30-member subsample, and perform subsampling after both the first month (Dec 2017) and second month (Jan 2018) of the forecasts to evaluate the sensitivity of the stratification to the lead time at which we do the subsampling. The stratification in each case is based on the daily Niño3.4 values and employs the KMEANS algorithm to separate the grand ensemble into L clusters, where we consider values of L = 3, 4, 5, 6, 7 and 10. We select, at random, a sample of members from each cluster (h), where the size of each sample is proportional to the size of corresponding population cluster ($\mathrm{n_h} = \mathrm{n} * \frac{\mathrm{N_h}}{\mathrm{N}}$, or the so-called proportional sampling). The monthly 30-member ensemble means obtained from such a sampling strategy are compared with those obtained by using a Monte Carlo approach and taking a random sample of 30 members from the population's 110 ensemble members.

The key metric we use to assess the value of stratification is the ratio ($\mathfrak{R}$) of the variance of the sample mean obtained from stratification $\bar{y_s}$ to that obtained from a simple random sampling of the population $\bar{y_r}$, namely:

$$\mathfrak{R} = \frac{Var(\bar{y_s})}{Var(\bar{y_r})} \tag{6}$$

An $\mathfrak{R}$ value below 1 indicates that the stratified sampling has a reduced variance of the mean relative to random sampling.

The ratio ($\mathfrak{R}$) shown in Fig. 8 for several different choices of the number of strata. The results illustrate that while there is a substantial reduction in $\mathfrak{R}$ when the sampling is performed during the first month (left panel), the benefit of the stratified sampling is rather quickly lost (after 3 months into the forecast). In contrast, when the sampling is done based on the second month (right panel), the benefit persists longer, extending into the 5th month. Based on these results, the GEOS-S2S-3 subsampling uses 4 strata and performs the subsampling after 2 months of forecast.

https://doi.org/10.5194/egusphere-2026-1340
Preprint. Discussion started: 6 May 2026

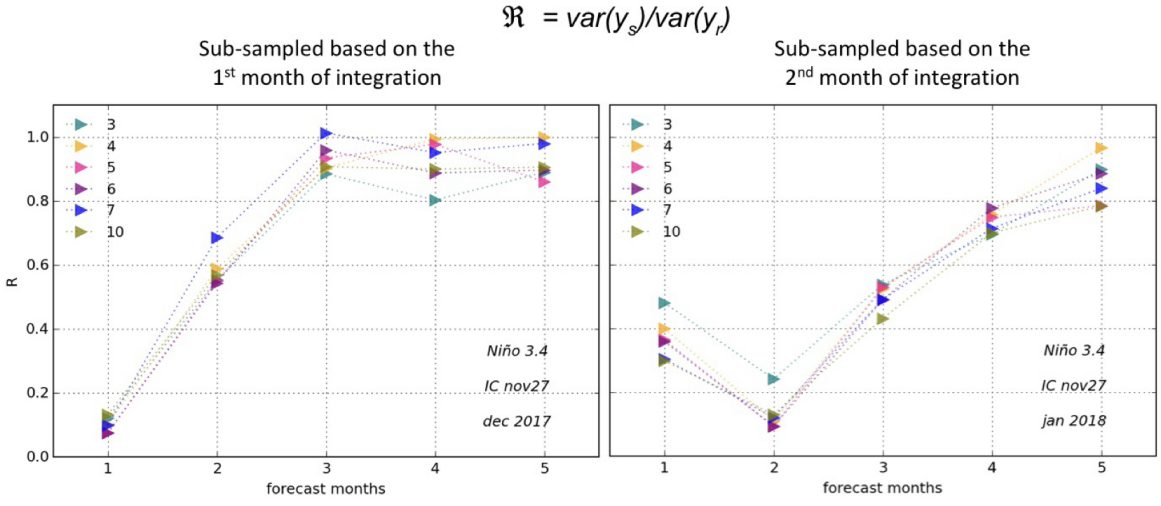

Figure 8. The ratio of variances of the ensemble mean of the Niño3.4 index for stratified versus random sampling, based on sampling after the first month of integration (Dec 2017, left panels) and after the second month of integration (Jan 2018, right panels). Results are presented for 3, 4, 5, 6, 7 and 10 clusters (strata). Variances are estimated from the Monte Carlo results.

4.2 Ensemble Spread Versus Forecast Uncertainty

The ensemble generation strategy implemented in GEOS S2S-3 was employed in generating a suite of retrospective forecasts for the period January 1991 to December, 2020. Based on the SMT and clustering methods described in Section 3, GEOS-S2S-3 seasonal forecasts now consist of a 40-member ensemble for the first three months, and sub-sample (Section 33.2) to a ten-member ensemble for the remaining six months of the forecast. Forecast are initialized every five days, and include one unperturbed member, four additional perturbed members (Section 33.1), and an additional ten perturbed members on the last initialization day of each month. Here we compare the ensemble spread to the forecast uncertainty for the Niño 3.4 index as a function start month and forecast lead time for the GEOS S2S-3 reforecasts and compare that to the results from GEOS S2S-2.

The key metric for evaluating the relationship between ensemble spread and forecast error (see the Appendix for the derivation) is:

$$R = \frac{SD_x}{SEE} \tag{7}$$

https://doi.org/10.5194/egusphere-2026-1340
Preprint. Discussion started: 6 May 2026

where SD_x is the mean intra-ensemble standard deviation of the forecasts (x), and SEE is the standard error of the estimate. The ratio R should ideally be equal to 1 (ensemble spread is equal to the forecast uncertainty). An under-dispersive model (the spread is lower than the error) has R values less than 1, while an over-dispersive model (spread is larger than the error) has R values greater than 1.

The ensemble spread characteristics from GEOS-S2S-2 and GEOS-S2S-3 (using the SMT and clustering method described in this manuscript) are shown in Figure 9 and exhibit distinctly different patterns across different forecast lead times and initialization periods in both model versions. GEOS S2S-2 shows R values close to 1 at shorter lead times (medium green), indicating that ensemble spread appropriately represents forecast uncertainty, is over-dispersive (light green to yellow to red) at long leads for forecasts initialized in boreal winter and early spring, and is under-dispersive (dark green to blue) at very short leads (months 1-2) when initialized during the May-June period. GEOS S2S-3 exhibits over-dispersive forecasts initialized during the boreal winter (NDJF) at lead months 1-5 and at short leads in early spring. GEOS-S2S-3 transitions to under-dispersive behavior at longer leads for forecasts initialized in boreal spring and at intermediate lead times (months 2-5) when initialized in boreal summer.

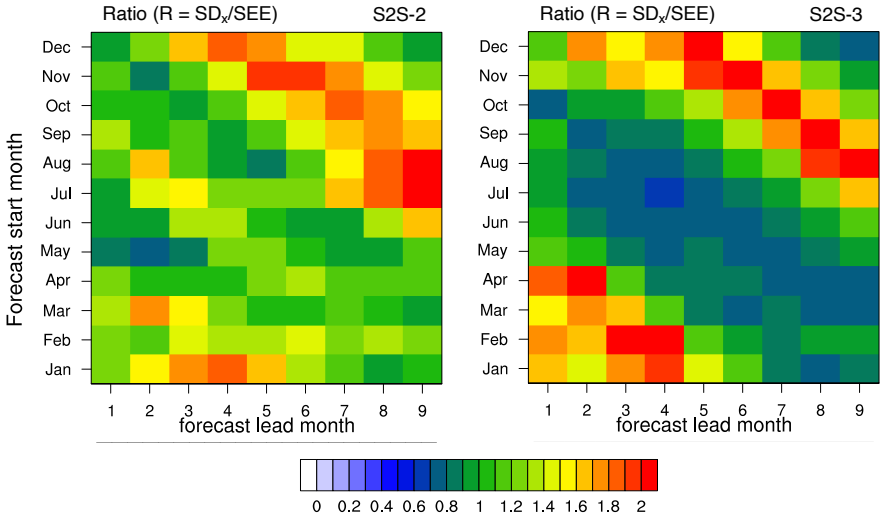

Figure 9. The ratio R of mean intra-ensemble standard deviation of forecasts (SD_x) with the standard error of the estimate (SEE) for the Niño 3.4 index as a function of start month and forecast lead month for the S2S-2 system (left) and the S2S-3 system (right). Results are based on four ensemble members in GEOS S2S-2, and 40 ensemble members for 1 and 2 month lead times and the 10 winners for lead times of 3 months and beyond in GEOS-S2S-3 for forecasts/hindcasts spanning the period 1991-2020.

The correlation skill of Niño 3.4 forecasts with respect to the Operational Sea Surface Temperature and Ice Analysis (OSTIA Donlon et al., 2012) observational estimates (not shown) demonstrates that both systems remain skillful at all leads (correlation

https://doi.org/10.5194/egusphere-2026-1340
Preprint. Discussion started: 6 May 2026

skill greater than 0.5), with the exception of forecasts initialized in January. GEOS S2S-3 is more skillful than GEOS S2S-2 at longer lead times when initialized in September through December, and less skillful at 2-5 month lead times for forecasts initialized in early spring. The differences in correlation skill are consistent with differences in the standard errors of the estimate (SEE, not shown) and the differences in the R-ratio, as the underdispersive nature of S2S-3 for forecasts initialized during spring and early summer (lead months 2-5) suggests that GEOS S2S-3 is less skillful as compared to GEOS S2S-2 for these initialization months and lead times. These results with high skill of GEOS S2S-3 initialized during early spring (at lead months 1-4) and winter are results are also reflected in standard errors of the estimate (SEE), with smaller values of SEE in GEOS S2S-3 than GEOS S2S-2 at these initialization months and lead times.

5 Summary and Conclusions

We have presented here a relatively simple to implement, yet effective, strategy for generating initial perturbations that are particularly relevant to the subseasonal to seasonal (S2S) prediction problem. We refer to that strategy as the Synchronized Multiple Time-lagged (SMT) approach. SMT uses the information about the temporal coherence of nearby coupled assimilation states to generate multiple perturbations that are imposed at a specified initial time, with pre-specified scaling factors determined as a fraction of the climatological variance.

We showed that by varying the time between the two analysis states from 1 to 10 days we are able to produce perturbations representative of a wide array of what appear to be physically realistic spatial structures in both the atmosphere and ocean. In fact, an examination of the leading eigenvectors (EOFs) of the relevant covariance matrices of the difference fields shows that the perturbations are in effect sampling well-known modes of variability that are relevant for S2S predictability. In the atmosphere these modes include middle latitude synoptic weather systems and teleconnections, tropical easterly waves, and MJO-like structures, while in the ocean these include various fluctuations in the thermocline and tropical Pacific instability waves. A further analysis of the EOFs associated with the difference fields showed that they are similar to the fastest growing modes of the linear propagator of the relevant dynamical system.

While we have yet to isolate the separate impacts of the different SMT perturbations on the overall growth of uncertainty (something that will require many more forecasts than are currently available), we are nevertheless convinced (given their connections to physically realistic modes of variability) that such perturbations should play an important role, especially in the growth of uncertainty in the atmosphere during the first month or so of the forecasts (the subseasonal forecasts).

Finally, anticipating continued computational resource limitations and timeliness constraints, we have outlined a strategy for reducing the forecast ensemble size after a specified lead time. Based on our initial results, the strategy, which involves performing a stratified sampling of the early larger ensemble in a way that accounts for the emerging directions of error growth, shows considerable promise for reducing the uncertainty in the ensemble mean (compared with simple random sampling) when the sampling is done after the second month of the forecasts. To some degree, the reduction in uncertainty appears to be maintained up to 3 months further into the forecasts.

https://doi.org/10.5194/egusphere-2026-1340
Preprint. Discussion started: 6 May 2026

Data availability. The current version of GEOS-S2S is available from the project website https://github.com/GEOS-ESM/GEOS-S2S-3/tree/main/src under the Apache License 2.0 (https://github.com/GEOS-ESM/GEOSgcm/blob/main/LICENSE). The version of the model used to produce the results used in this paper is archived on repository under DOI https://doi.org/10.5281/zenodo.19009940, as are input data and scripts to run the model and produce the plots for all the simulations presented in this paper (DOI https://doi.org/10.5281/zenodo.19009727). The file specification document that describes the available output from GEOS-S2S and the other experimental forecasts is available from: https://gmao.gsfc.nasa.gov/pubs/docs/Nakada1033.pdf.

Appendix A: R ratio of forecast spread to uncertainty

Following Barnston et al. (2015) (see also Schubert et al., 2019) we compare the mean intra-ensemble standard deviation of the forecasts (x):

$$SD_x = \sqrt{\overline{\langle (x - \langle x \rangle)^2}} \tag{A1}$$

where the overbar indicates a long term mean over all the years of forecasts/hindcasts (1991-2023 for GEOS-S2S-2), the angle brackets denote an ensemble mean and the standard error of the estimate (SEE, the standard error of a simple linear regression in which the predictor is the ensemble mean forecast) written as:

$$SEE = SD_y * \sqrt{1 - cor^2_{y\langle x \rangle}} \tag{A2}$$

here $cor^2_{y\langle x \rangle}$ is the correlation between the ensemble mean forecast ($\langle x \rangle$) and the observations (y), given by:

$$cor_{y\langle x \rangle} = \frac{\overline{(y - \bar{y}) * (\langle x \rangle - \overline{\langle x \rangle})}}{SD_y * SD_{\langle x \rangle}} \tag{A3}$$

where the standard deviation of the observations (y) is:

$$SD_y = \sqrt{\overline{(y - \bar{y})^2}} \tag{A4}$$

and that of the ensemble forecasts $\langle x \rangle$ is:

$$SD_y = \sqrt{\overline{(\langle x \rangle - \overline{\langle x \rangle})^2}} \tag{A5}$$

We can then define the ratio:

$$R = \frac{SD_x}{SEE} \tag{A6}$$

which should ideally be equal to 1 (ensemble spread is equal to the forecast uncertainty). An under-dispersive model has R values less than 1, while an over-dispersive model has R values greater than 1.

Author contributions. Author AB conducted the experiments, performed the analysis and produced the majority of the figures. Author SS conceived of the SMT method and connected it to singular vector decomposition. Author AM conceived of the subsampling strategy and

https://doi.org/10.5194/egusphere-2026-1340
Preprint. Discussion started: 6 May 2026

created the schematic. Authors YKL and PY performed the analysis of the results from GEOS-S2S-3, and PY produced the figures showing those results

Competing interests. There are no competing interests

Acknowledgements. All the authors carried out the experimental design and computations at NASA's Goddard Space Flight Center. The high performance computing resources are from the NASA Center for Climate Simulate (NCCS) and the NASA Advanced Supercomputing (NAS) Division of the Ames Research Center. This work towards the development of GEOS-S2S-3 was funded under the NASA MAP-supported GMAO 'core' funding.

https://doi.org/10.5194/egusphere-2026-1340
Preprint. Discussion started: 6 May 2026

References

Adler, R., Huffman, G., Chang, A., Ferraro, R., Xie, P., Janowiak, J., Rudolf, B., Schneider, U., Curtis, S., Bolvin, D., Gruber, A., and Susskind, J.: The Version2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979 Present), Journal of Hydrometeorology - J HYDROMETEOROL, 4, https://doi.org/10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2, 2003.

Akella, S. R. and Suarez, M.: The Atmosphere-Ocean Interface Layer of NASA's Goddard Earth Observing System Model and Data Assimilation System Volume 51, Tech. rep., The National Aeronautics and Space Administration, 2018.

Baehr, J. and Piontek, R.: Ensemble initialization of the oceanic component of a coupled model through bred vectors at seasonal-to-interannual timescales, Geoscientific Model Development, 7, 453–461, https://doi.org/10.5194/gmd-7-453-2014, 2014.

Barnston, A. G., Tippett, M. K., van den Dool, H. M., and Unger, D. A.: Toward an Improved Multimodel ENSO Prediction, Journal of Applied Meteorology and Climatology, 54, 1579–1595, https://doi.org/10.1175/JAMC-D-14-0188.1, 2015.

Berner, J., Fossell, K. R., Ha, S.-Y., Hacker, J. P., and Snyder, C.: Increasing the skill of probabilistic forecasts: Understanding performance improvements from model-error representations, Mon. Wea. Rev., 143, 1295–1320, 2015.

Bloom, S. C., Takacs, L. L., da Silva, A. M., and Ledvina, D.: Data Assimilation Using Incremental Analysis Updates, Monthly Weather Review, 124, 1256–1271, 1996.

Borovikov, A.: GEOS-S2S forecast output, https://doi.org/10.5281/zenodo.19009727, 2018.

Borovikov, A., Cullather, R., Kovach, R., Marshak, J., Vernieres, G., Vikhliaev, Y., Zhao, B., and Li, Z.: GEOS-5 Seasonal Forecast System, Climate Dynamics, https://doi.org/10.1007/s00382-017-3835-2, 2017.

Buchard, V., Randles, C. A., da Silva, A. M., Darmenov, A., Colarco, P. R., Govindaraju, R., Ferrare, R., Hair, J., Beyersdorf, A. J., Ziemba, L. D., and Yu, H.: The MERRA-2 Aerosol Reanalysis, 1980 Onward. Part II: Evaluation and Case Studies, Journal of Climate, 30, 6851 – 6872, https://doi.org/10.1175/JCLI-D-16-0613.1, 2017.

Buizza, R. and Palmer, T. N.: The singular-vector structure of the atmospheric general circulation, Journal of the Atmospheric Sciences, 52, 1434–1456, https://doi.org/10.1175/1520-0469(1995)052<1434:TSVSOT>2.0.CO;2, 1995.

Chin, M., Ginoux, P., Kinne, S., Torres, O., Holben, B. N., Duncan, B. N., Martin, R. V., Logan, J. A., Higurashi, A., and Nakajima, T.: Tropospheric Aerosol Optical Thickness from the GOCART Model and Comparisons with Satellite and Sun Photometer Measurements, J. Atmos. Sci., 59, 461–483, 2002.

Christensen, H. M., Moroz, I. M., and Palmer, T. N.: Stochastic and Perturbed Parameter Representations of Model Uncertainty in Convection Parameterization, Journal of the Atmospheric Sciences, 72, 2525–2540, https://doi.org/10.1175/JAS-D-14-0250.1, 2015.

Cochran, W. G.: Sampling Techniques, John Wiley & Sons, Inc., New York, 2nd edn., 1963.

Colarco, P., da Silva, A., Chin, M., and Diehl, T.: Online simulations of global aerosol distributions in the NASA GEOS-4 model and comparisons to satellite and ground-based aerosol optical depth, J. Geophys. Res. Atmos., 115, 2010.

Dalcher, A., Kalnay, E., and Hoffman, R. N.: Medium range lagged average forecasts, Mon. Wea. Rev., 116, 402–416, https://doi.org/10.1175/1520-0493(1988)116<0402:MRLAF>2.0.CO;2, 1988.

Delsole, T., Trenary, L., and Tippett, M.: The Weighted-Average Lagged Ensemble, Journal of Advances in Modeling Earth Systems, 9, https://doi.org/10.1002/2017MS001128, 2017.

Donlon, C. J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E., and Wimmer, W.: The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system, Remote Sensing of Environment, 116, 140 – 158, https://doi.org/http://dx.doi.org/10.1016/j.rse.2010.10.017, advanced Along Track Scanning Radiometer(AATSR) Special Issue, 2012.

https://doi.org/10.5194/egusphere-2026-1340
Preprint. Discussion started: 6 May 2026

Feldstein, S. B.: The Timescale, Power Spectra, and Climate Noise Properties of Teleconnection Patterns, Journal of Climate, 13, 4430–4440, 2000.

Griffies, S. M.: Elements of the Modular Ocean Model (MOM) (2012 Release), Tech. rep., Geophysical Fluid Dynamics Laboratory, Princeton, NJ, 2012.

Griffies, S. M., Gnanadesikan, A., Dixon, K. W., Dunne, J. P., Gerdes, R., Harrison, M. J., Rosati, A., Russell, J. L., Samuels, B. L., Spelman, M. J., Winton, M., and Zhang, R.: Formulation of an ocean model for global climate simulations, Ocean Science, 1, 45–79, 2005.

Ham, Y.-G.: Optimal Initial Perturbations for Ensemble Prediction of the Madden-Julian Oscillation during Boreal Winter, Climate Dynamics, 25, 14, https://doi.org/10.1007/s00382-012-1419-4, 2012.

Hill, C., DeLuca, C., Balaji, Suarez, M., and Da Silva, A.: The Architecture of the Earth System Modeling Framework, Computing in Science & Engineering, 6, 18–28, 2004.

Hoffman, R. N. and Kalnay, E.: Lagged average forecasting, an alternative to Monte Carlo forecasting, Tellus A, 35A, 100–118, https://doi.org/https://doi.org/10.1111/j.1600-0870.1983.tb00189.x, 1983.

Hunke, E. and Lipscomb, W.: CICE, The Los Alamos Sea Ice Model, Version 4.1, 2008.

Imada, Y. and Kimoto, M.: Parameterization of Tropical Instability Waves and Examination of Their Impact on ENSO Characteristics, Journal of Climate, 25, 4568 – 4581, https://doi.org/10.1175/JCLI-D-11-00233.1, 2012.

Kalnay, E.: Atmospheric Modeling, Data Assimilation and Predictability, Cambridge University Press, Cambridge, 2003.

Koster, R. D., Suarez, M. J., Ducharne, A., Stieglitz, M., and Kumar, P.: A Catchment-based Approach to Modeling Land Surface Processes in a General Circulation Model: 1. Model Structure, Journal of Geophysical Research: Atmosphere, 105, 24 809–24 822, 2000.

Magnusson, L., Leutbecher, M., and Källén, E.: Comparison between Singular Vectors and Breeding Vectors as Initial Perturbations for the ECMWF Ensemble Prediction System, Monthly Weather Review - MON WEATHER REV, 136, 4092–4104, https://doi.org/10.1175/2008MWR2498.1, 2008.

Manrique-Suñén, A., Palma, L., Gonzalez-Reviriego, N., Doblas-Reyes, F. J., and Soret, A.: Subseasonal predictions for climate services, a recipe for operational implementation, Climate Services, 30, 100 359, https://doi.org/https://doi.org/10.1016/j.cliser.2023.100359, 2023.

Molod, A., Takacs, L., Suarez, M., and Bacmeister, J.: Development of the GEOS-5 atmospheric general circulation model: evolution from MERRA to MERRA2, Geoscientific Model Development, 8, 1339–1356, https://doi.org/10.5194/gmd-8-1339-2015, 2015.

Molod, A., Hackert, E., Vikhliaev, Y., Zhao, B., Barahona, D., Vernieres, G., Borovikov, A., Kovach, R. M., Marshak, J., Schubert, S., Li, Z., Lim, Y.-K., Andrews, L. C., Cullather, R., Koster, R., Achuthavarier, D., Carton, J., Coy, L., Friere, J. L. M., Longo, K. M., Nakada, K., and Pawson, S.: GEOS-S2S Version 2: The GMAO High-Resolution Coupled Model and Assimilation System for Seasonal Prediction, Journal of Geophysical Research: Atmospheres, 125, e2019JD031 767, https://doi.org/10.1029/2019JD031767, e2019JD031767 10.1029/2019JD031767, 2020.

Orbe, C., Oman, L. D., Strahan, S. E., Waugh, D. W., Pawson, S., Takacs, L. L., and Molod, A. M.: Large-Scale Atmospheric Transport in GEOS Replay Simulations, Journal of Advances in Modeling Earth Systems, 9, 2545–2560, https://doi.org/https://doi.org/10.1002/2017MS001053, 2017.

Penny, S. G., Kalnay, E., Carton, J. A., Hunt, B. R., Ide, K., Miyoshi, T., and Chepurin, G. A.: The Local Ensemble Transform Kalman Filter and the Running-in-place Algorithm Applied to a Global Ocean General Circulation Model, Nonlinear Processes in Geophysics, 20, 1031–1046, 2013.

Reichle, R. H. and Liu, Q.: Observation-Corrected Precipitation Estimates in GEOS-5, Technical Report Series on Global Modeling and Data Assimilation, 35, 101, 2014.

https://doi.org/10.5194/egusphere-2026-1340
Preprint. Discussion started: 6 May 2026

Rienecker, M. M., Suarez, M. M. J., Todling, R., Bacmeister, J., Tackacs, L., Liu, H.-C., Gu, W., Sienkiewicz, M., Koster, R. R. D., Gelaro, R., Stajner, I., Nielsen, J. E., Takacs, L., Liu, H.-C., Gu, W., Sienkiewicz, M., Koster, R. R. D., Gelaro, R., Stajner, I., and Nielsen, J. E.: The GEOS-5 Data Assimilation System — Documentation of Versions 5.0.1, 5.1.0, and 5.2.0, Tech. rep., NASA Goddard Space Flight Center, Greenbelt, MD, 2008.

Sansom, P., Ferro, C., Stephenson, D., Goddard, L., and Mason, S.: Best Practices for Postprocessing Ensemble Climate Forecasts. Part I: Selecting Appropriate Recalibration Methods, Journal of Climate, 29, https://doi.org/10.1175/JCLI-D-15-0868.1, 2016.

Scaife, A. A. and Smith, D. M.: A signal-to-noise paradox in climate science, npj Climate and Atmospheric Science, 1, 28, 2018.

Schubert, S., Suarez, M., Schemm, J.-K., and Epstein, E.: Dynamically Stratified Monte Carlo Forecasting, Monthly Weather Review, 120, 1077 – 1088, https://doi.org/10.1175/1520-0493(1992)120<1077:DSMCF>2.0.CO;2, 1992.

Schubert, S., Borovikov, A., Lim, Y.-K., and Molod, A.: Ensemble Generation Strategies Employed in the GMAO GEOS-S2S Forecast System, Tech. rep., NASA/TM-2019-104606, Vol. 53, https://ntrs.nasa.gov/citations/20200001232, 2019.

Shinoda, T., Lyman, J. M., Johnson, G. C., Lavín, M. F., and Kawai, Y.: Statistical representation of equatorial waves and tropical instability waves in the Pacific Ocean, Atmospheric Research, 94, 37–44, https://doi.org/10.1016/j.atmosres.2008.06.010, 2009.

Spaeth, H.: Cluster Analysis Algorithms for Data Reduction and Classification of Objects, E. Horwood, Chichester, Eng. ; New York : Halsted Press, 1980.

Strang, G.: Linear Algebra and Its Applications, Wellesley-Cambridge Press, Wellesley, MA, book Description: Gilbert Strang's textbooks have changed the entire approach to learning linear algebra – away from abstract vector spaces to specific examples of the four fundamental subspaces: the column space and nullspace of A and A'., 2009.

Strobach, E., Molod, A., Trayanov, A., Forget, G., Campin, J.-M., Hill, C., and Menemenlis, D.: Three-to-Six-Day Air–Sea Oscillation in Models and Observations, Geophysical Research Letters, 47, e2019GL085 837, https://doi.org/10.1029/2019GL085837, e2019GL085837 10.1029/2019GL085837, 2020.

Suarez, M., Trayanov, A., Hill, C., Schopf, P., and Vikhliaev, Y.: MAPL: A High-level Programming Paradigm to Support More Rapid and Robust Encoding of Hierarchical Trees of Interacting High-performance Components, in: Proceedings of the 2007 Symposium on Component and Framework Technology in High-performance and Scientific Computing - CompFrame '07, pp. 11–20, ACM Press, New York, New York, USA, 2007.

Takacs, L. L., Suarez, M. J., and Todling, R.: The Stability of Incremental Analysis Update, Monthly Weather Review, 146, 3259 – 3275, https://doi.org/10.1175/MWR-D-18-0117.1, 2018.

Toth, Z. and Kalnay, E.: Ensemble Forecasting at NCEP and the Breeding Method, Monthly Weather Review, 125, 3297 – 3319, https://doi.org/10.1175/1520-0493(1997)125<3297:EFANAT>2.0.CO;2, 1997.

Vitart, F. and Takaya, Y.: Lagged ensembles in sub-seasonal predictions, Quarterly Journal of the Royal Meteorological Society, 147, https://doi.org/10.1002/qj.4125, 2021.

Wallace, J. M. and Gutzler, D. S.: Teleconnections in the Geopotential Height Field during the Northern Hemisphere Winter, Monthly Weather Review, 109, 784–812, https://doi.org/10.1175/1520-0493(1981)109<0784:TITGHF>2.0.CO;2, 1981.

Xie, P. and Arkin, P. A.: Global Precipitation: A 17-Year Monthly Analysis Based on Gauge Observations, Satellite Estimates, and Numerical Model Outputs, Bulletin of the American Meteorological Society, 78, 2539 – 2558, https://doi.org/10.1175/1520-0477(1997)078<2539:GPAYMA>2.0.CO;2, 1997.

Yang, S.-C., Cai, M., Kalnay, E., Rienecker, M., Yuan, G., and Toth, Z.: ENSO Bred Vectors in Coupled Ocean-Atmosphere General Circulation Models, Journal of Climate, 19, https://doi.org/10.1175/JCLI3696.1, 2004.