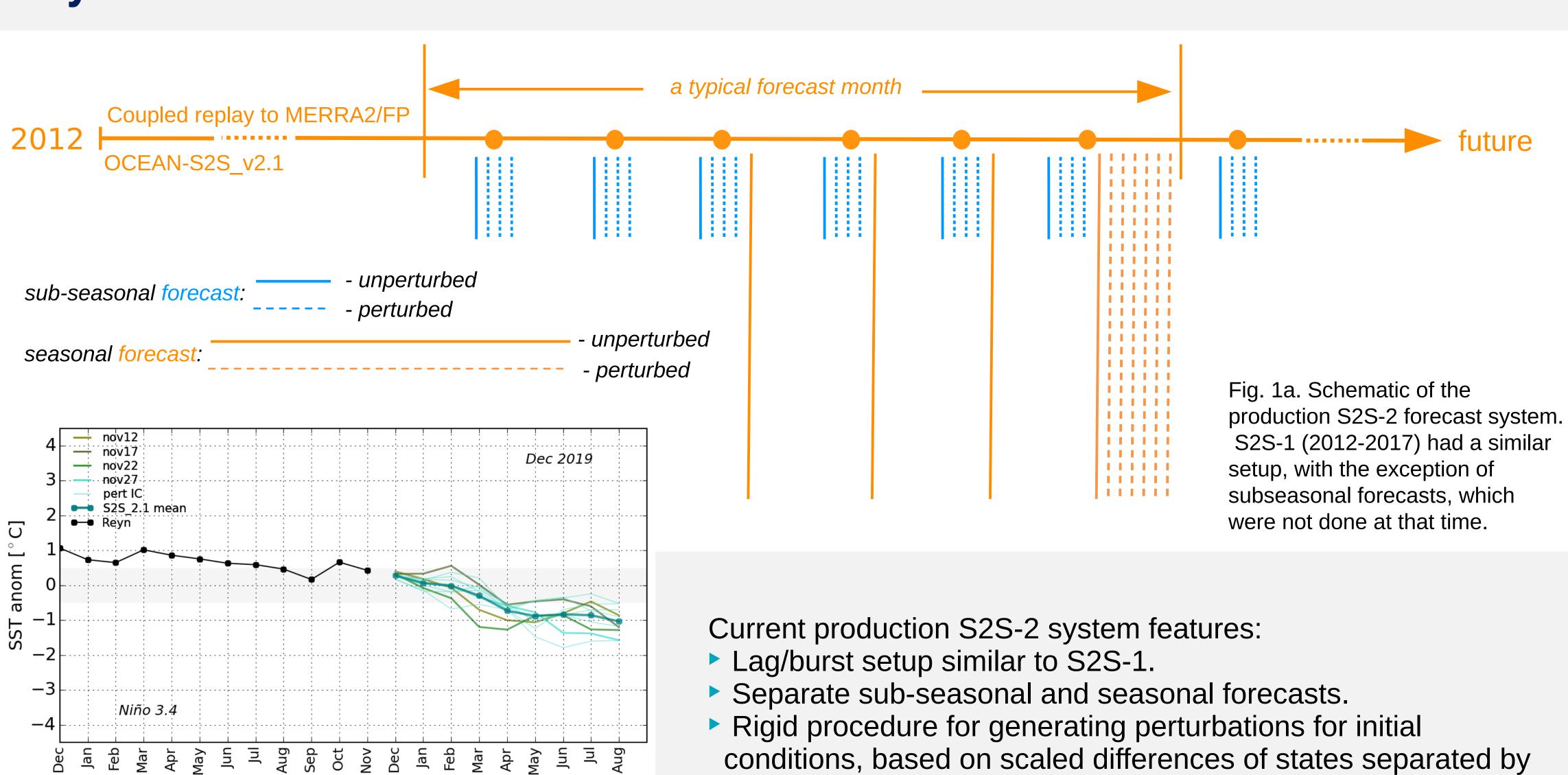


Global Modeling & Assimilation Office

Why do we need new ensembles for Subseasonal-to-Seasonal forecasts?



Learning from S2S-1 and S2S-2

Is the ensemble spread an indicator of forecast uncertainty?

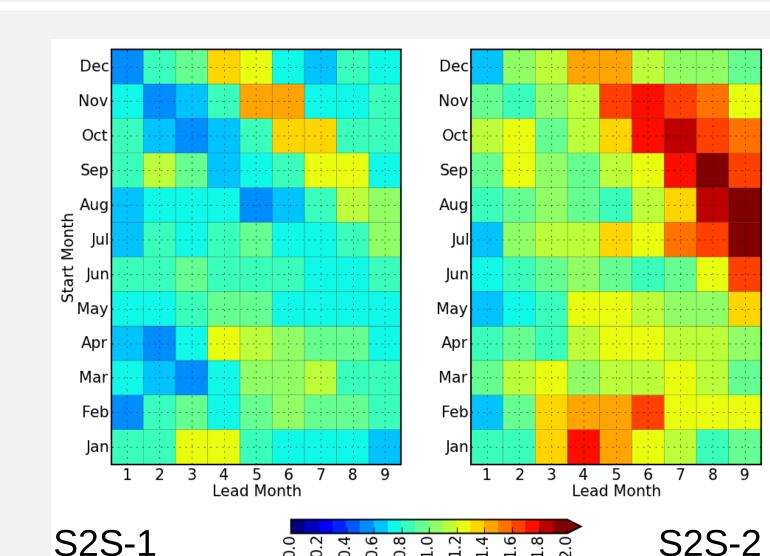
Let SDy be the standard deviation of the observation (y), cor_{yy}^2 the squared correlation between the ensemble mean forecast (x) and the observation, σ the standard deviation of the intra-ensemble spread, then $SEE = SDy \sqrt{1 - cor_{yy}^2}$ and R = 1 σ/SEE , which should be close to 1 for a perfect model:

if R < 1 the model is under dispersive

Fig. 1b. An example of the production S2S-2 Nino3.4 plume.

if R > 1 model is over dispersive

Fig. 2. R for both forecast system versions for for Niño3.4 SST, all initial months, all leads.



Ensemble design for S2S-3

Motivation:

- For ENSO improve the under-dispersion at short lead time; control the over-dispersion at long leads.
- For sub-seasonal teleconnections improve the ensemble mean skill by increasing ensemble size.

Explore ideas for GEOS S2S-3:

- The various patterns (eigenvectors) of perturbations and scaling.
- Different combinations of lag and burst.
- Use large ensemble for season-long forecasts, then sub-sample and continue with fewer members for the long-range forecasts.

Methods:

- Bursts of forecasts on a single date using initial conditions perturbations, generated from instantaneous states from coupled analysis at varying separations. Synchronized Multiple Time-lagged (SMT) approach.
- Stratified sampling to select ensemble members for long range forecast.

Spatial patterns of perturbations. Span all scales.

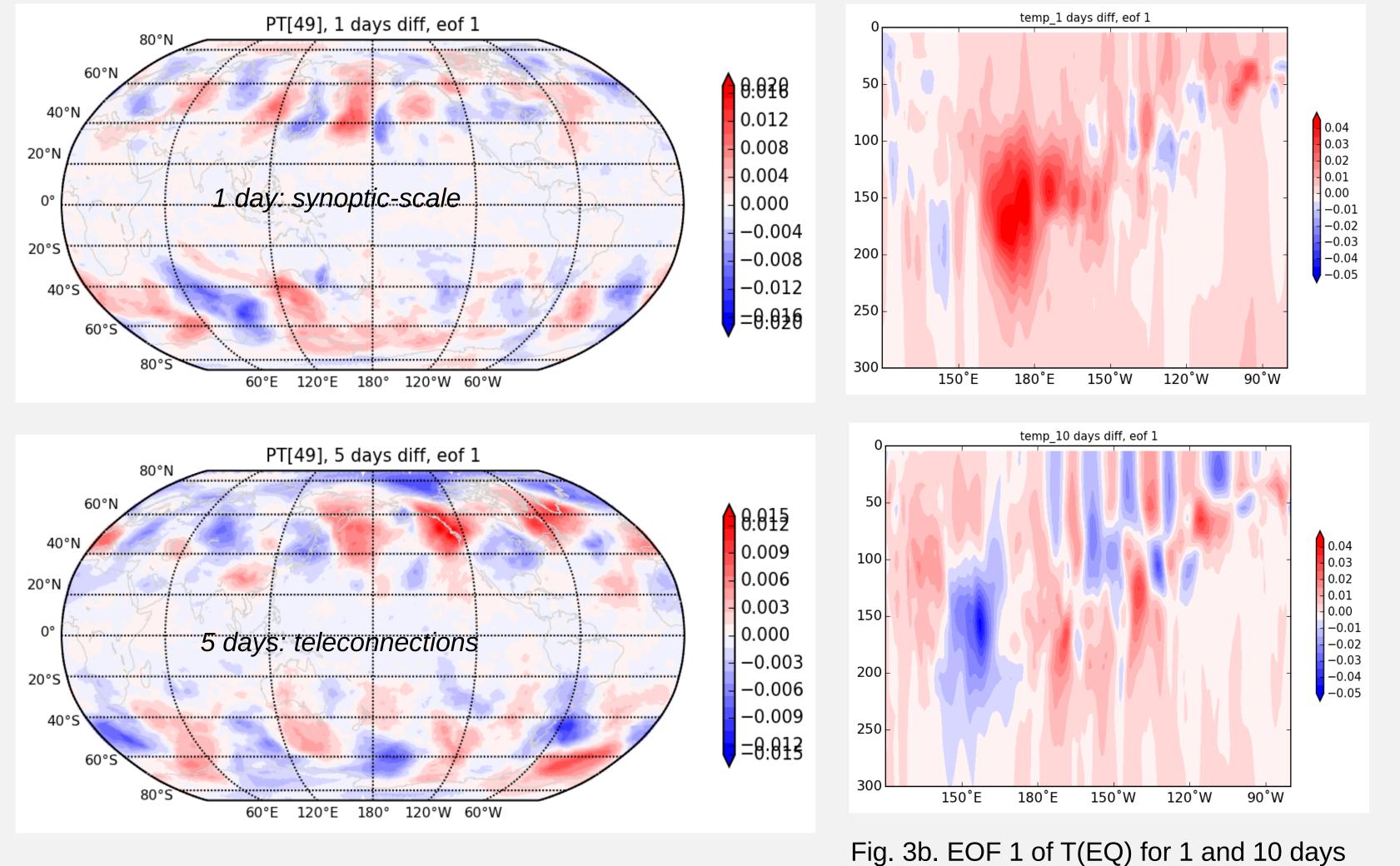


Fig.3a. EOF 1 of Potential Temperature (500mb) for 1 and 5 days

By varying the separation time between nearby analysis states we are able to generate a wide array of different types of atmospheric and oceanic perturbations that represent physically realistic and important modes of variability.

Scaling. To make the perturbations amplitude a small fraction (10% in terms of STD) of the natural variability and independent of the separation, we produced an average set of scaling factors that vary only with season and states separation for ocean and atmosphere variables.

Stratified sampling. KMEANS.

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 that we capture the leading directions (in phase space) of error growth. This can be especially important when the ensemble is characterized by more than one dominant direction of error growth.

The population of size *N* is divided into L disjoint strata, where $n_{\rm b}$ (N_b) are the number of members of the sample (population) in stratum *h*. Then each stratum is sampled in proportion to its representation in the population.

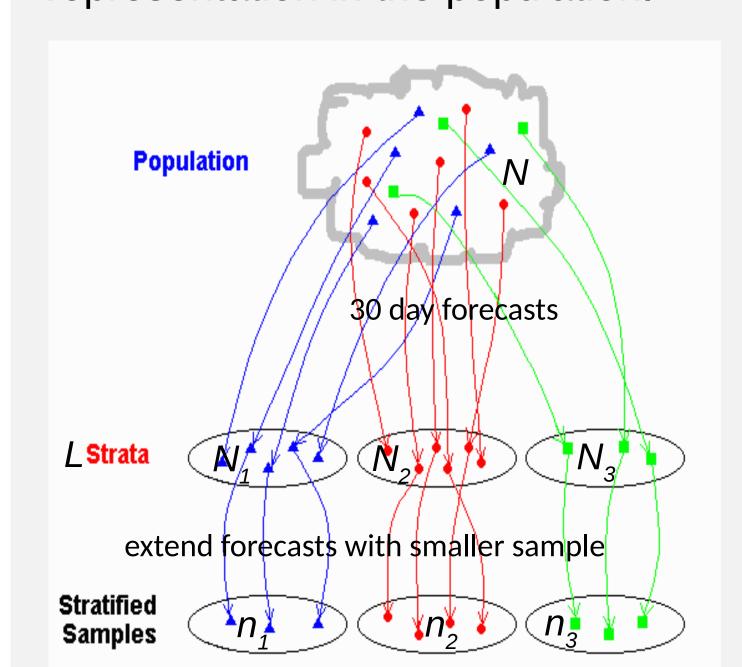


Fig. 5a. Schematic illustrating clustering procedure. Here the number of clusters is L=3 with N_1 , N_2 and N_3 their respective populations sampled down to n_1 , n_2 and n_3 .

ensemble is well spanned.

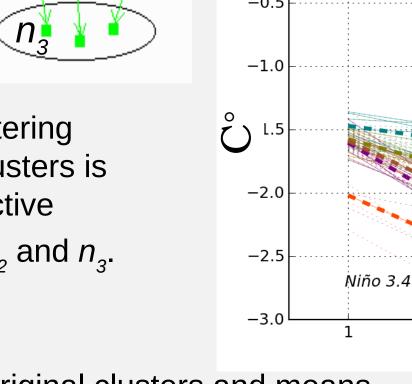
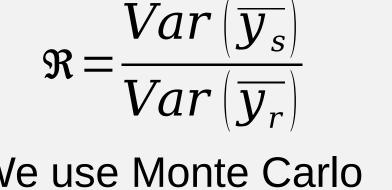


Fig. 5b. Sub-sampling EXAMPLE. Top: Daily Niño 3.4 index values, original clusters and means. Middle: Daily sub-sampled clusters and means (dotted lines) and original means (dashed lines). Bottom: Monthly values, extended forecast, solid lines are the sub-samples ensemble members, dashed – all the original members, thick dashed lines – cluster means. The envelope of the original

Quantifying the results.

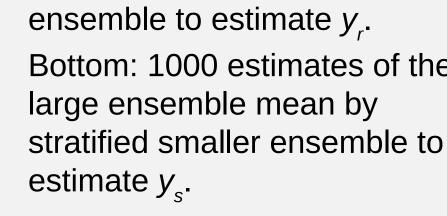
Performing the stratification very early in the forecasts emphasizes the variance structure of the initial perturbations, and those structures are not well maintained as the forecasts evolve beyond the first

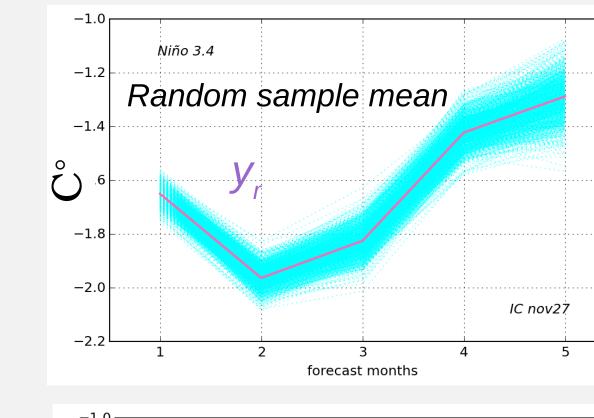
By the second month the clusters are more likely to reflect the uncertainties associated with the underlying dynamical evolution of the climate system, which are maintained much longer into the forecast.



We use Monte Carlo approach (with 1000 random seeds) to estimate the ensemble means and then the variances of these

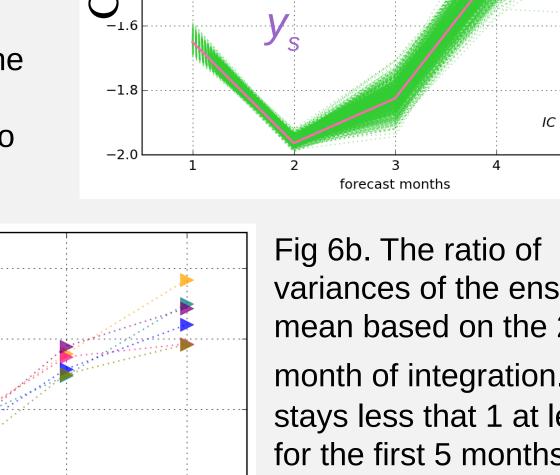
the large ensemble mean by a randomly sub-sampled smaller Bottom: 1000 estimates of the





Stratified sample mean

Fig. 6a. Top: 1000 estimates of



variances of the ensemble mean based on the 2nd month of integration. \Re stays less that 1 at least for the first 5 months of integration. Results are presented for 3,4,5,6,7 and 10 clusters (strata) We find little benefit from increasing the number of strata beyond 4 or 5.

Testing various types and combinations of lag/burst.

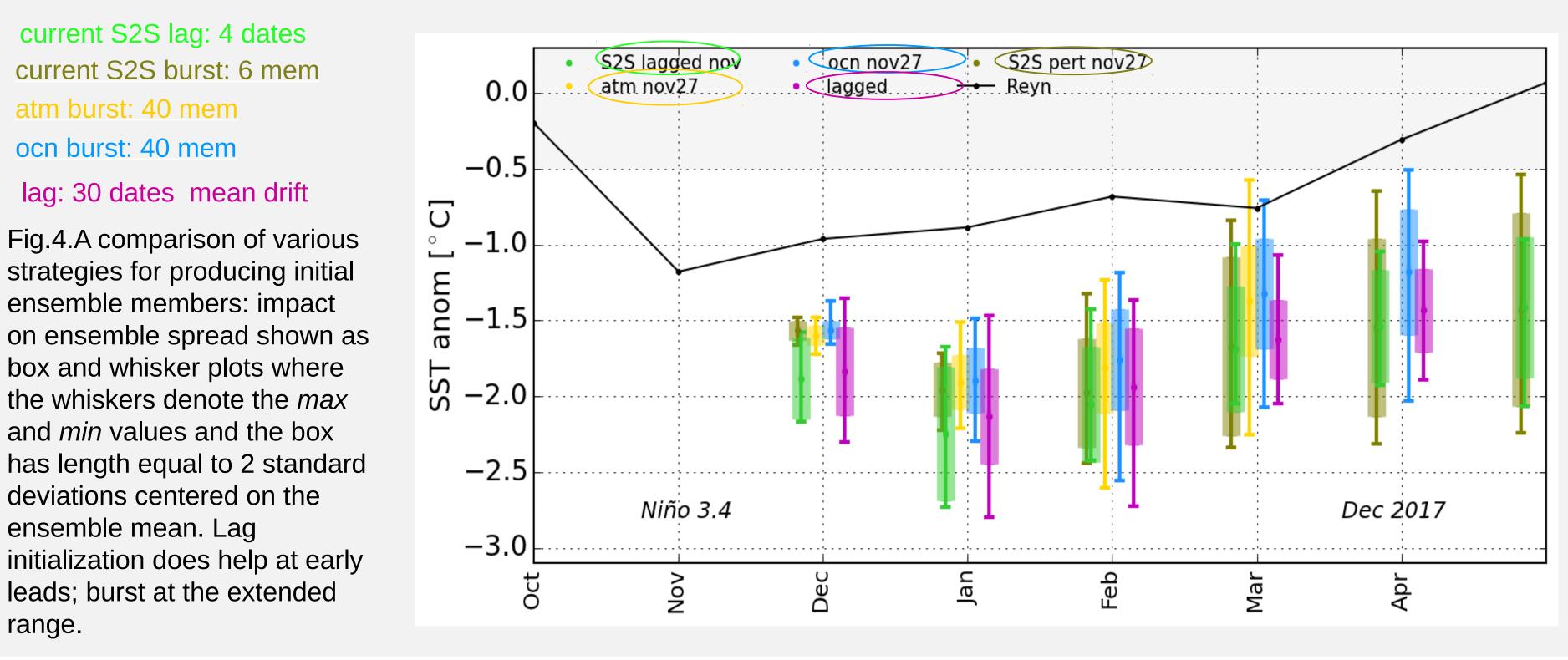
current S2S lag: 4 dates current S2S burst: 6 mem atm burst: 40 mem ocn burst: 40 mem

lag: 30 dates mean drift

Fig.4.A comparison of various strategies for producing initial ensemble members: impact on ensemble spread shown as box and whisker plots where the whiskers denote the *max* and *min* values and the box has length equal to 2 standard deviations centered on the ensemble mean. Lag

leads; burst at the extended

range.



perturbations.

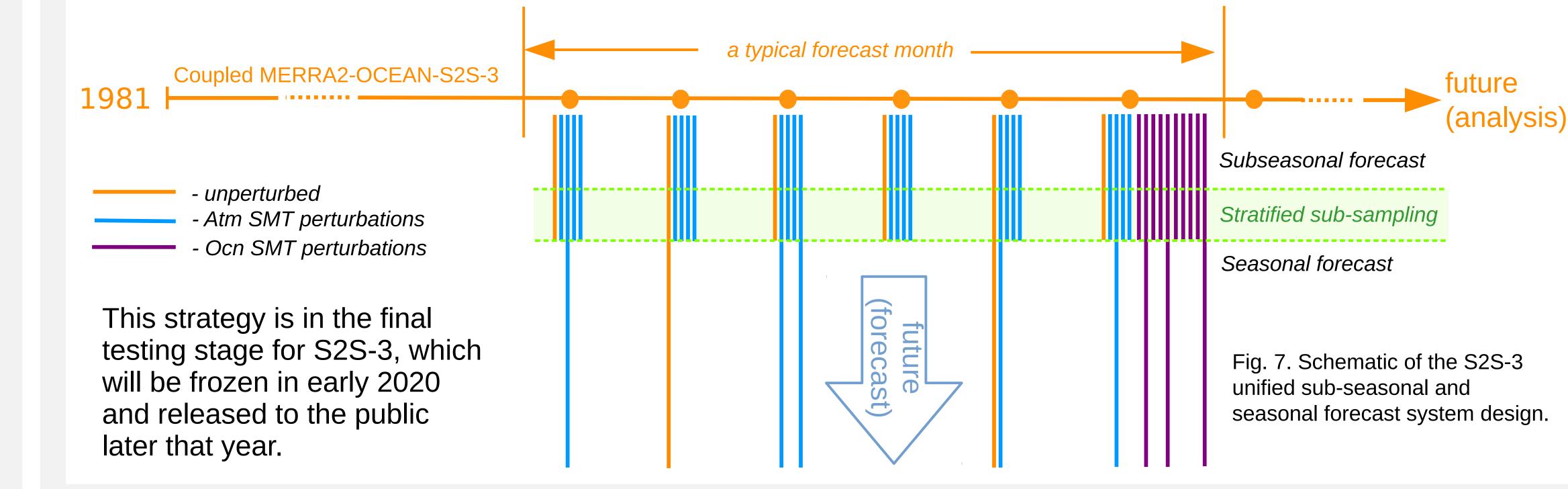
separation.1 day: variations in the

wave-type variability.

thermocline. 10 days: vertically-coherent

Fig. 3c. Scaling factors for atmospheric

Final ensemble design



5(1) days for seasonal(subseasonal) forecasts.