SMALL-SCALE VARIABILITY IN SST: ESTIMATION AND USES IN CLIMATE ANALYSES

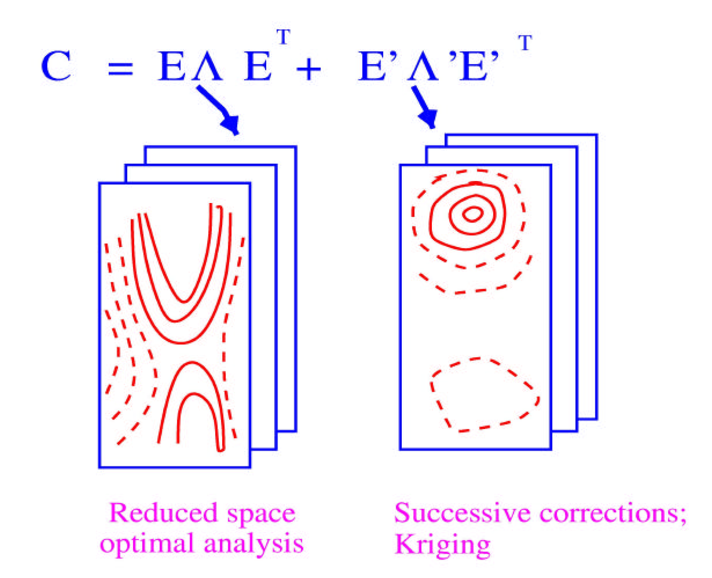
Alexey Kaplan, **Mark Cane,**
alexeyk@ldeo.columbia.edu **mcane@ldeo.columbia.edu**
Climate Modeling Group, Division of Ocean and Climate Physics
Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY 10964, United States

1. MOTIVATION

The principal goal of this project is to use sea surface temperature (SST) data from the MODIS instrument to improve the estimate of global fields of SST over the past two centuries. Climate change studies require long historical data sets to put the present and future changes in the context of natural variability, and to verify climate models. Studies of interannual, decadal, and longer climate variations require timeseries long enough to encompass many realizations. Since raw historical data does not provide consistent coverage over large areas, investigators use analyzed gridded fields, such as the well-known SST products of Reynolds and Smith, Kaplan, or the Hadley Center (GISST). All of these analyses products require estimates of observational error, including the sampling error arising because the few observations available within a given grid box and time period are not representative of the average over that grid box and period. This sampling error is a consequence of sub-grid scale variations, here referred to as small-scale variability (SSV). The almost 200 years of SST data in the historical archive does not have the spatial and temporal resolution to estimate SSV globally. But the few years (3 years at present) of high resolution MODIS data are adequate to provide what is needed for sampling error estimates.

2. FORMALISM OF REDUCED SPACE OBJECTIVE ANALYSES

APPROXIMATING COVARIANCE



GENERAL PROBLEM OF RECONCILING MODELS WITH DATA

$$\mathcal{T}_{n+1} = A_n\mathcal{T}_n + \varepsilon_n^{\rm m}, \quad n = 1,\ldots,N-1$$
$$\mathcal{T}_n^o = H_n\mathcal{T}_n + \varepsilon_n^o, \quad n = 1,\ldots,N.$$

$$\langle\varepsilon_n^o\rangle = 0, \qquad \langle\varepsilon_n^o\varepsilon_n^{o\,T}\rangle = R_n, \quad n = 1,\ldots,N$$
$$\langle\varepsilon_{n_1}^o\varepsilon_{n_2}^{o\,T}\rangle = 0, \quad n_1 \neq n_2, \quad \langle\varepsilon_{n_1}^o\mathcal{T}_{n_2}^T\rangle = 0, \quad n_1,n_2 = 1,\ldots,N,$$
$$\langle\varepsilon_n^{\rm m}\rangle = 0, \qquad \langle\varepsilon_n^{\rm m}\varepsilon_n^{{\rm m}\,T}\rangle = Q_n, \quad n = 1,\ldots,N-1$$
$$\langle\varepsilon_{n_1}^{\rm m}\varepsilon_{n_2}^{{\rm m}\,T}\rangle = 0, \quad n_1 \neq n_2, \quad \langle\varepsilon_{n_1}^{\rm m}\mathcal{T}_{n_2}^T\rangle = 0, \quad n_1,n_2 = 1,\ldots,N-1$$
$$\langle\varepsilon_{n_1}^o\varepsilon_{n_2}^{{\rm m}\,T}\rangle = 0, \quad n_1 = 1,\ldots,N, \qquad n_2 = 1,\ldots,N-1$$

MINIMIZATION OF THE FULL COST FUNCTION:

$$\mathbf{S}[\mathcal{T}_1,\mathcal{T}_2,\ldots,\mathcal{T}_N] = \Sigma_{n=1}^N(H_n\mathcal{T}_n - \mathcal{T}_n^o)^T R_n^{-1}(H_n\mathcal{T}_n - \mathcal{T}_n^o) + \Sigma_{n=1}^{N-1}(\mathcal{T}_{n+1} - A_n\mathcal{T}_n)^T Q_n^{-1}(\mathcal{T}_{n+1} - A_n\mathcal{T}_n)$$

OPTIMAL SMOOTHER (OS) and KALMAN FILTER (KF)

"Sweep up" – **KF**:

$$\hat{\mathcal{T}}_n^a = \hat{\mathcal{T}}_n^f + K_n\left(\hat{\mathcal{T}}_n^o - H_n\hat{\mathcal{T}}_n^f\right),$$
$$\hat{\mathcal{T}}_n^f = A_n\hat{\mathcal{T}}_{n-1}^a,$$
$$K_n = P_n^f H_n^T\left(H_n P_n^f H_n^T + R_n\right)^{-1}$$
$$P_n^a = (I_M - K_nH_n)P_n^f$$
$$P_n^f = A_{n-1}P_{n-1}^a A_{n-1}^T + Q_{n-1}, \qquad n = 2,3,\ldots,N$$

"Sweep down" – **OS**:

$$\hat{\mathcal{T}}_n^s = \hat{\mathcal{T}}_n^a + G_n\left(\hat{\mathcal{T}}_{n+1}^s - A_n\hat{\mathcal{T}}_n^a\right), \qquad G_n = P_n^a A_n^T (P_{n+1}^f)^{-1},$$
$$P_n^s = P_n^a + G_n\left(P_{n+1}^s - P_{n+1}^f\right)G_n^T, \qquad n = N-1,\ldots,2,1$$

Classical data assimilation and objective analyses techniques require dealing with covariance matrices (*P*,*Q*,*R*) whose order is equal to the dimension of the system space. Reduced space versions carry error covariance calculations in the low-dimensional space of the predetermined set of patterns (*E*) for which we take a set (<100) of leading empirical orthogonal functions from a model run.

SPACE REDUCTION

$$C = E\Lambda E^T + E'\Lambda' E'^T$$
$$\mathcal{T}_n = E\alpha_n + \varepsilon_n^r, \quad n = 1,\ldots,N$$

ESTIMATION PROBLEM IN THE REDUCED SPACE

$$\mathcal{T}_n^o = H_nE\alpha_n + (H_n\varepsilon_n^r + \varepsilon_n^o) \stackrel{\rm def}{=} \mathcal{H}_n\alpha_n + \check{\varepsilon}_n^o, \quad n = 1,\ldots,N,$$

$$\alpha_{n+1} = \mathcal{A}_n\alpha_n + E^T\varepsilon_n^{\rm m} \stackrel{\rm def}{=} \mathcal{A}_n\alpha_n + \check{\varepsilon}_n^{\rm m}, \quad n = 1,\ldots,N-1.$$

$$\mathcal{Q}_n = \langle\check{\varepsilon}_n^{\rm m}\check{\varepsilon}_n^{{\rm m}\,T}\rangle = E^T\langle\varepsilon_n^{\rm m}\varepsilon_n^{{\rm m}\,T}\rangle E = E^TQ_nE$$

$$\mathcal{R}_n = \langle\check{\varepsilon}_n^o\check{\varepsilon}_n^{o\,T}\rangle = \langle(H_n\varepsilon_n^r + \varepsilon_n^o)(H_n\varepsilon_n^r + \varepsilon_n^o)^T\rangle = \langle\varepsilon_n^o\varepsilon_n^{o\,T}\rangle + H_n\langle\varepsilon_n^r\varepsilon_n^{r\,T}\rangle H_n^T \stackrel{\rm def}{=} R_n + H_nQ^rH_n^T \stackrel{\rm def}{=} R_n + R_n'.$$

REDUCED SPACE OPTIMAL ANALYSIS

Cost function:

$$\mathcal{S}[\alpha_1,\alpha_2,\ldots,\alpha_N] = \Sigma_{n=1}^N(\mathcal{H}\alpha_n - \mathcal{T}_n^o)^T\mathcal{R}_n^{-1}(\mathcal{H}\alpha_n - \mathcal{T}_n^o) + \Sigma_{n=1}^{N-1}(\alpha_{n+1} - \mathcal{A}_n\alpha_n)^T\mathcal{Q}_n^{-1}(\alpha_{n+1} - \mathcal{A}_n\alpha_n).$$

KF:

$$\alpha_n^a = \alpha_n^f + \mathcal{K}_n\left(\mathcal{T}_n^o - \mathcal{H}_n\alpha_n^f\right),$$
$$\alpha_n^f = \mathcal{A}_n\alpha_{n-1}^a,$$
$$\mathcal{K}_n = \left(\mathcal{H}_n^T\mathcal{R}_n^{-1}\mathcal{H}_n + \mathcal{P}_n^{f\,-1}\right)^{-1}\mathcal{H}_n^T\mathcal{R}_n^{-1}$$
$$\mathcal{P}_n^a = (I_L - \mathcal{K}_n\mathcal{H}_n)\mathcal{P}_n^f$$
$$\mathcal{P}_n^f = \mathcal{A}_{n-1}\mathcal{P}_{n-1}^a\mathcal{A}_{n-1}^T + \mathcal{Q}_{n-1}, \qquad n = 2,3,\ldots,N$$

OS:

$$\alpha_n^s = \alpha_n^a + G_n\left(\alpha_{n+1}^s - \mathcal{A}_n\alpha_n^a\right),$$
$$G_n = \mathcal{P}_n^a\mathcal{A}_n^T(\mathcal{P}_{n+1}^f)^{-1},$$
$$\mathcal{P}_n^s = \mathcal{P}_n^a + G_n\left(\mathcal{P}_{n+1}^s - \mathcal{P}_{n+1}^f\right)G_n^T, \qquad n = N-1,\ldots,2,1$$

3. TECHNIQUE VALIDATION

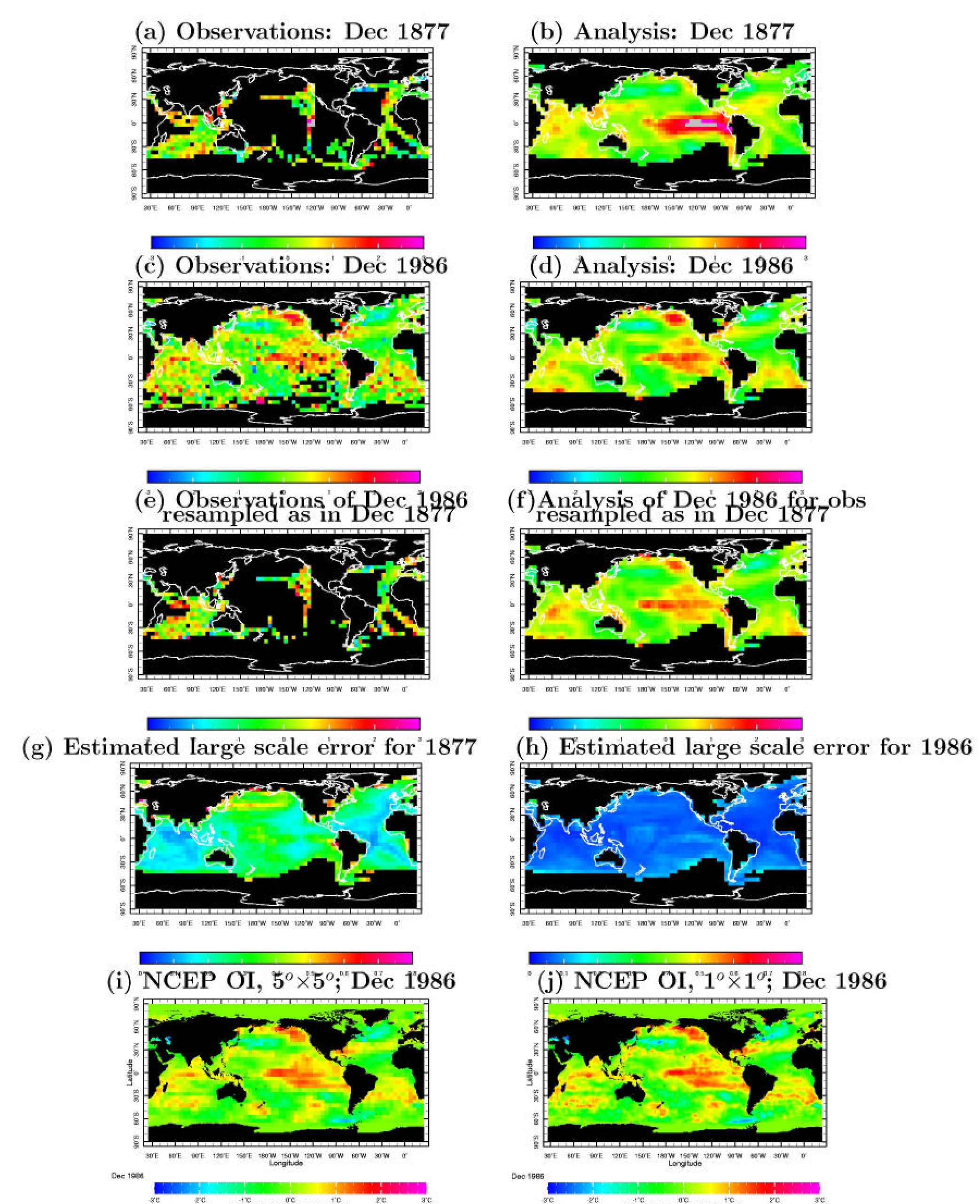
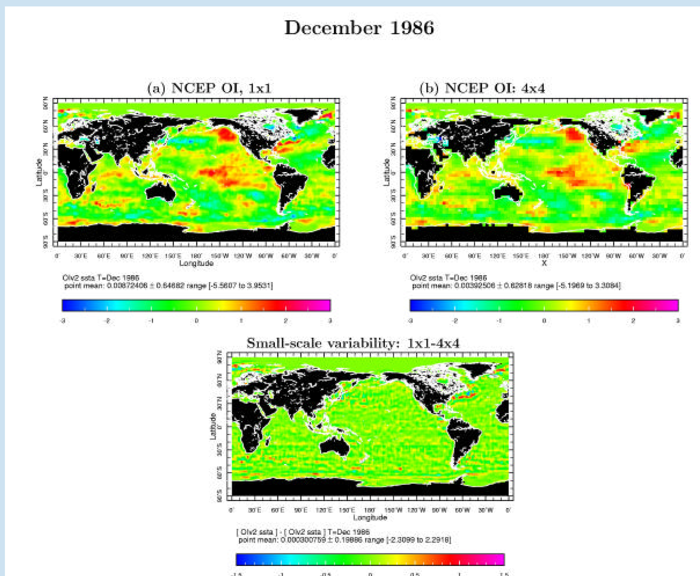


Figure 2: Available SST observations and their RS OS analysis for December 1877 (panels (a) and (b)) with verification through the experiment with 1986 data: simulated OS analysis for December 1986 using the data distribution of 1877 (panels (e) and (f)) versus the standard OS analysis for December 1986 with all available data (panels (c) and (d)). Also shown are large scale errors in the two reconstructions (panels (e) and (f)) and the NCEP OI December 1986 field presented in (i) $5^o\times5^o$ and (j) $1^o\times1^o$ resolution. Units are ^{o}C.

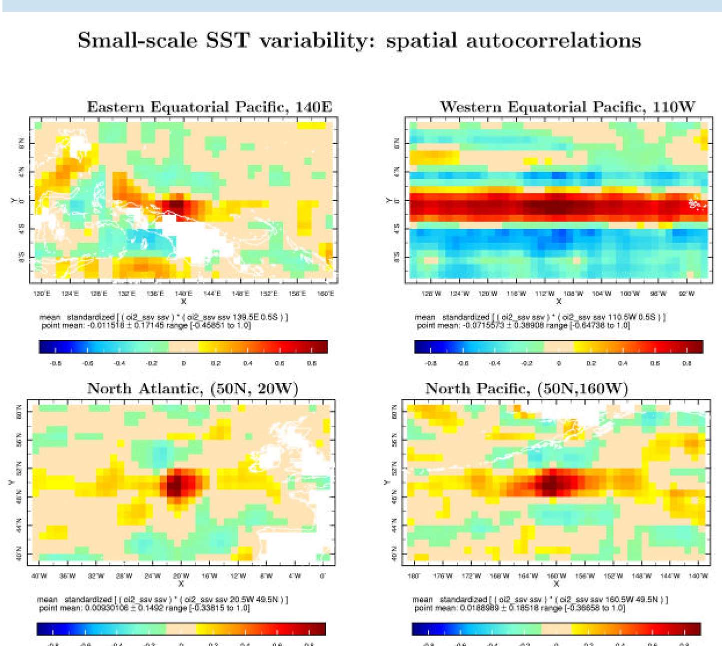
There is a lot of room for improvement , as far as reconstruction of small-scale variability is concerned.
We need to know its statistics.

4. SMALL-SCALE VARIABILITY

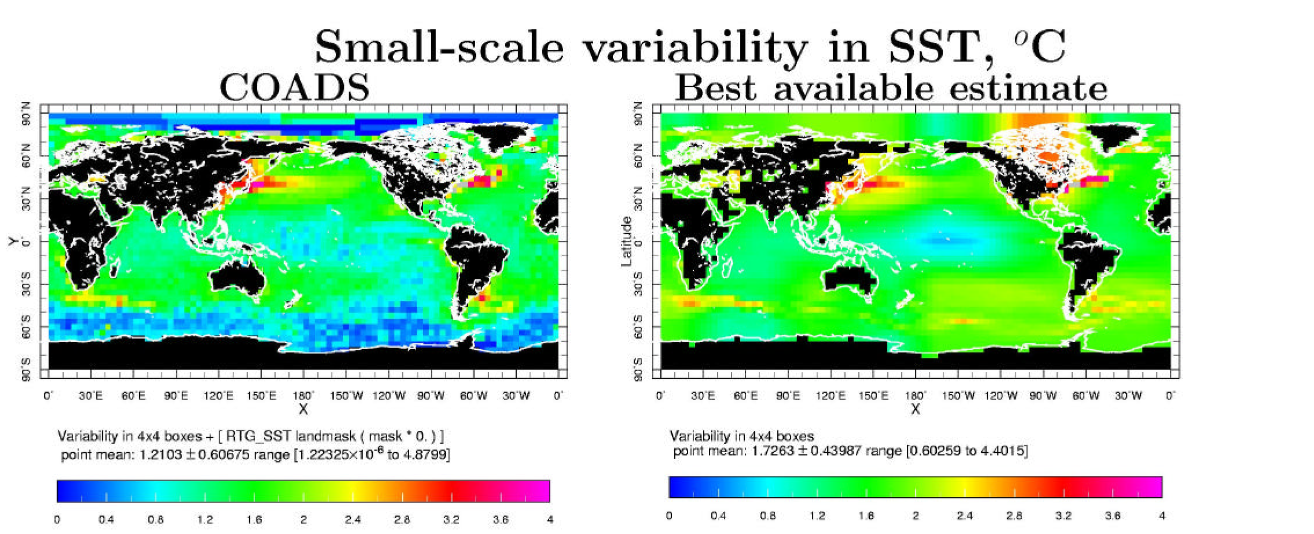


Clearly non-stationary covariance structure!

5. AUTOCORRELATION STRUCTURES IN SMALL-SCALE VARIABILITY



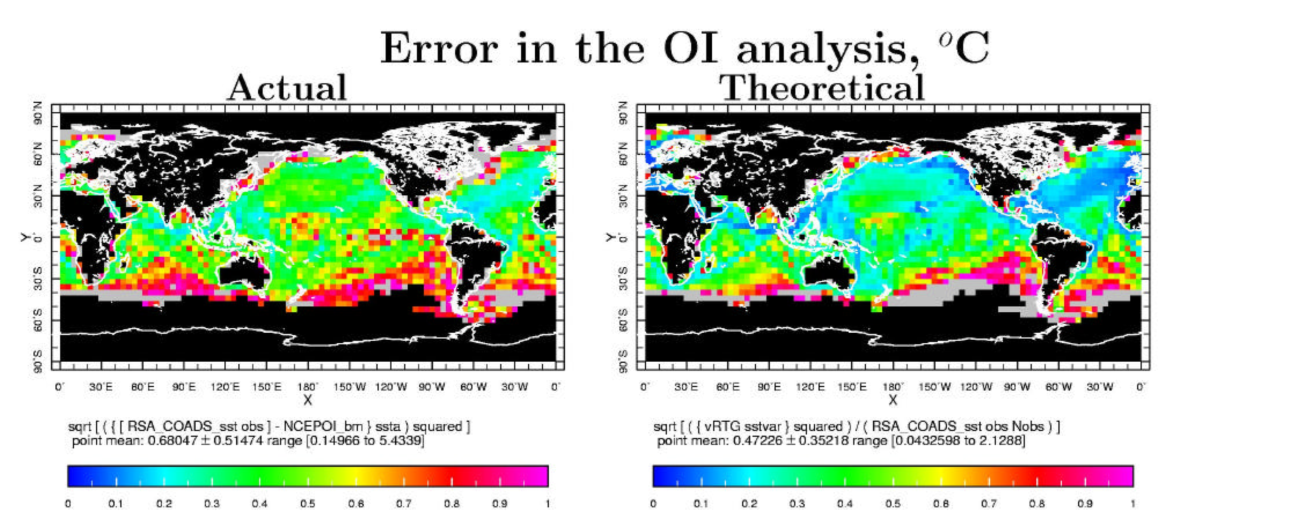
6. AVAILABLE ESTIMATES OF SMALL-SCALE VARIABILITY



Estimates of standard deviation in SSV-SST, ^{o}C, for $4^o\times4^o$ monthly bins. Shown are: (left) estimate from in situ data (COADS) only; (right) best available estimate – see text for explanations.

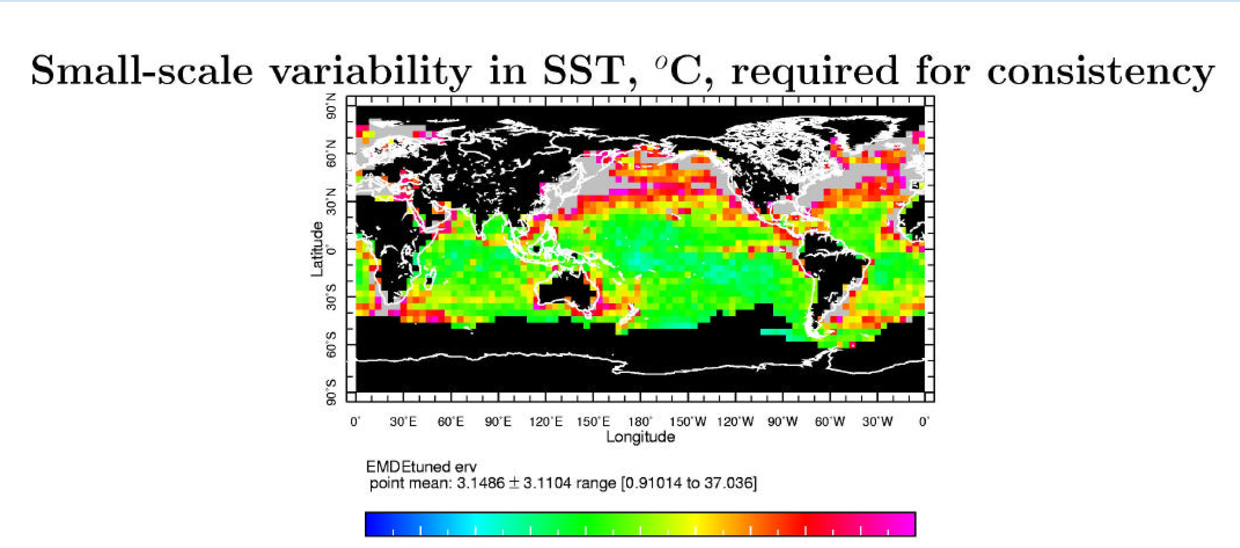
Existing estimates of small-scale variability in the SST are clearly inadequate. This results in incorrect analysis error estimates and suboptimal analyses too.

7. DISCREPANCY IN THE ANALYSES



Estimates of standard deviation in the OI analysis of in situ (COADS) data , ^{o}C, for $4^o\times4^o$ monthly bins. Shown are: (left) actual error estimated as RMS difference between the OI analysis of COADS data and the benchmark data set; (right) theoretical error of the OI analysis.

8. CONJECTURED VARIABILITY



Estimates of standard deviation in SSV-SST, ^{o}C, needed for consistency with COADS–benchmark difference (backed out from formula (1)).

9. PLANNED TASKS

1. Maps of SSV-SST inside bins of sizes varying from 0.25 to 5 degree spatially and from a week to a month temporally will be produced. Systematic comparison of MODIS data with other satellite mission products (AVHRR, ATSR) and high-resolution model analyses will be performed. This analysis will be repeated for both skin and bulk SST products, in order to evaluate possible influences of the skin-to-bulk conversion algorithms on our results; possible cloud contamination and the effects of diurnal warming will be investigated as well.

2. The total spacextime SSV-SST will be separated into spatial and temporal components. Estimates of the latter from MODIS will be compared with estimates from TOGA-TAO buoys.

3. Results obtained in earlier parts of the project will be summarized in an isotropic approximation: dependence of the mean SSV on the size of the spacextime bin will be presented as a function of location. Conversely, these dependences will be evaluated by a modified variogram technique, and results compared.

4. More sophisticated descriptions of the SSV via anisotropic approximation of covariance will be attempted as well. We will characterize local autocorrelation maps in various locations and will study a possible connection of large-scale and small-scale variability in SST.

5. We'll apply the filtered spectral analysis of to MODIS data. We will attempt to produce frequency and 2D location-dependent wavenumber spectra of sea surface temperatures.

6. Finally, the consistency and utility of our findings will receive a field test: we will redo the OI analysis of the 20 years of COADS data using improved estimates of the SSV-SST from MODIS. We will then redo our analysis of historical SSTs.