Use of Satellite Data for Gridded SST Analyses of Pre-Satellite Period

Alexey Kaplan

Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY 10964, USA

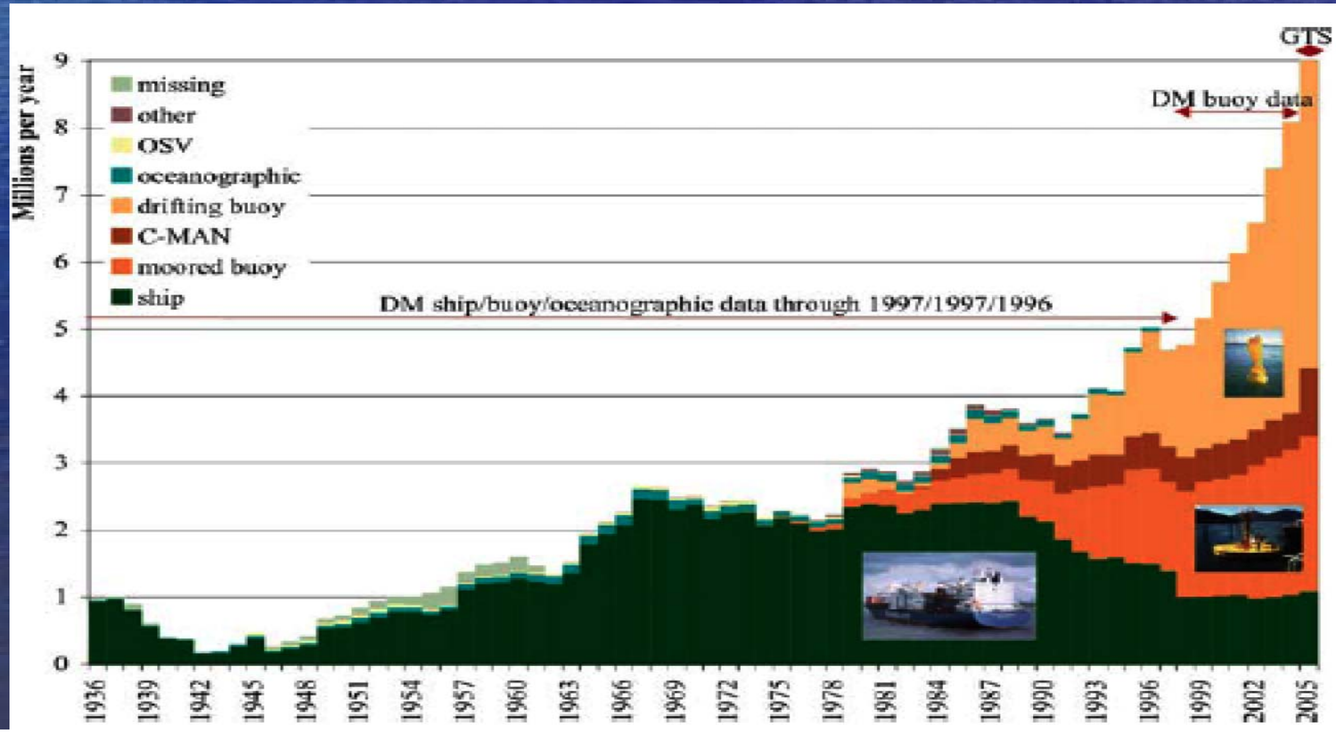
Alicia Karspeck

National Center for Atmospheric Research, Boulder, CO 80307, USA

OUTLINE

- Historical data availability
- Optimal Interpolation (OI) as gridding and analysis method
- Estimation of input covariances (for observational and background error)
- Utility of the space reduction for coarse gridding of historical in situ data
- Importance of satellite data for estimating parameters of higher resolution OI
- Representing uncertainty by samples from the posterior distribution

Transition to the modern Ocean Observing System

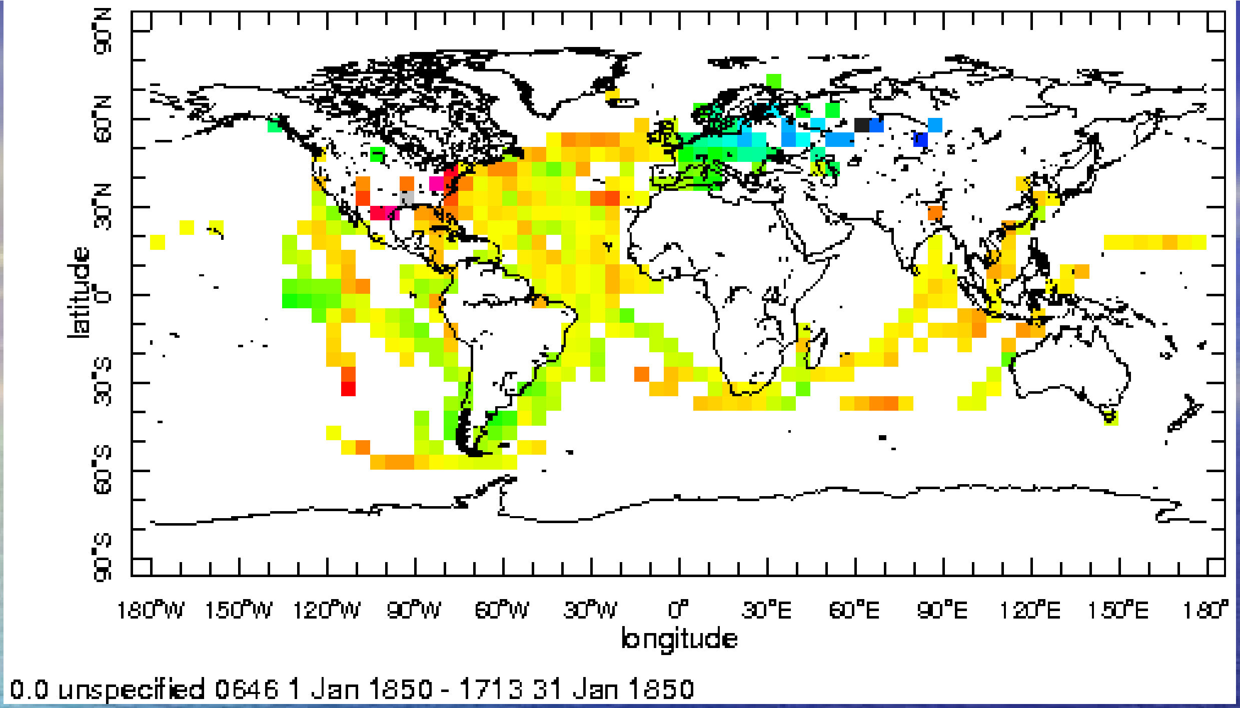


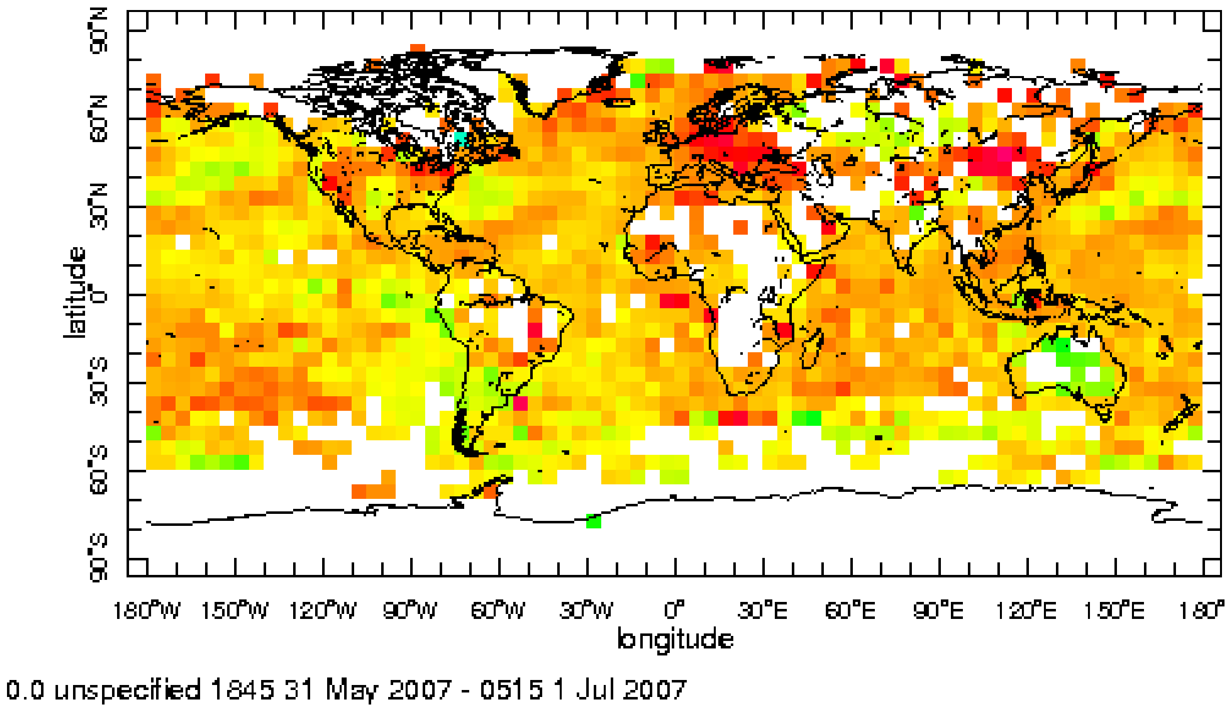
From *Woodruff et al.* [2008],
In *Climate Variability and Extremes during the Past 100 Years*, **Bronniman et al. (eds.)**

Jan 1850

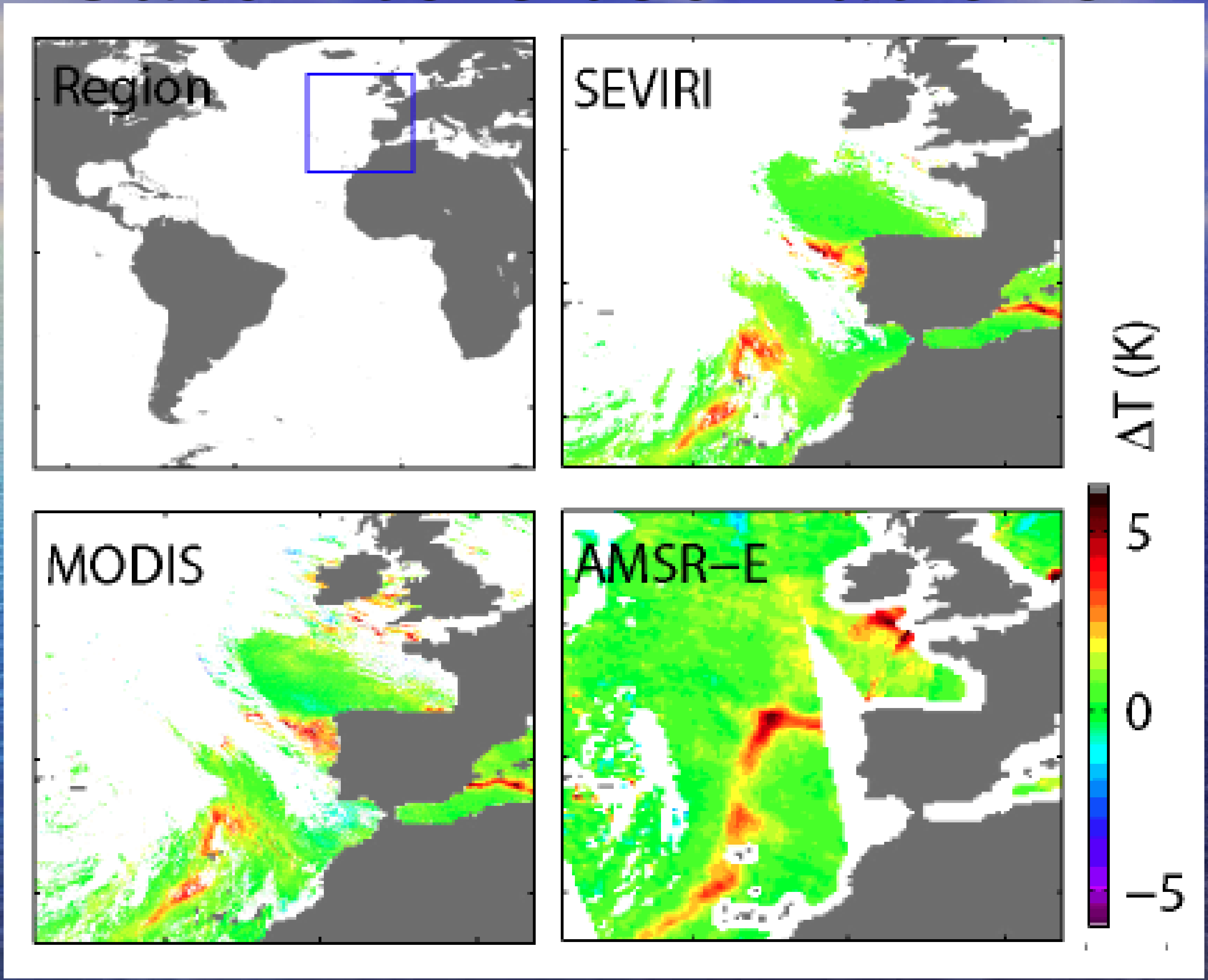
In situ data:
HadSST2
[Rayner et al., 2006]

Jul 2007





Satellite Observations



Donlon et al. [2010], OceanObs'09, Community White Paper

PROBLEM

Richness and volume of these data sets notwithstanding, as descriptors of detailed historical SST variability they are quite incomplete, affected by large errors, or else of rather short time coverage.

Yet many applications require SST fields interpolated onto a regular grid, with no spatial or temporal gaps, and uncertainty estimates.

Optimal Interpolation (OI)

$$T = T_B + e_B$$

$$HT = T_o + e_o$$

$$< e_B > = < e_o > = < e_B e_o^T > = 0$$

$< e_B e_B^T > = C$ ← bckgr err (signal) cov

$< e_o e_o^T > = R$ ← observ err covariance

Solution minimizes the cost function

$$S[T] = (HT - T_o)^T R^{-1} (HT - T_o) + (T - T_B)^T C^{-1} (T - T_B)$$

$$T_{OI} = P_{OI} (H^T R^{-1} T_o + C^{-1} T_B), \quad P_{OI} = (H^T R^{-1} H + C^{-1})^{-1}$$

What we infer from the data, given the OI assumptions

$$T \sim \mathcal{N}(T_{OI}, P_{OI}),$$

where $T_{OI} = P_{OI}(H^T R^{-1} T_o + C^{-1} T_B)$

$$= T_B + CH^T(HCH^T + R)^{-1}(HT_o - T_B),$$

$$P_{OI} = (H^T R^{-1} H + C^{-1})^{-1}$$

$$= C - CH^T(HCH^T + R)^{-1} HC$$

Matrix dimensions are R: $N_{obs} x N_{obs}$, C: $N_g x N_g$,

i.e., for 5° grid N_g ~2000, for 1° grid N_g ~50,000

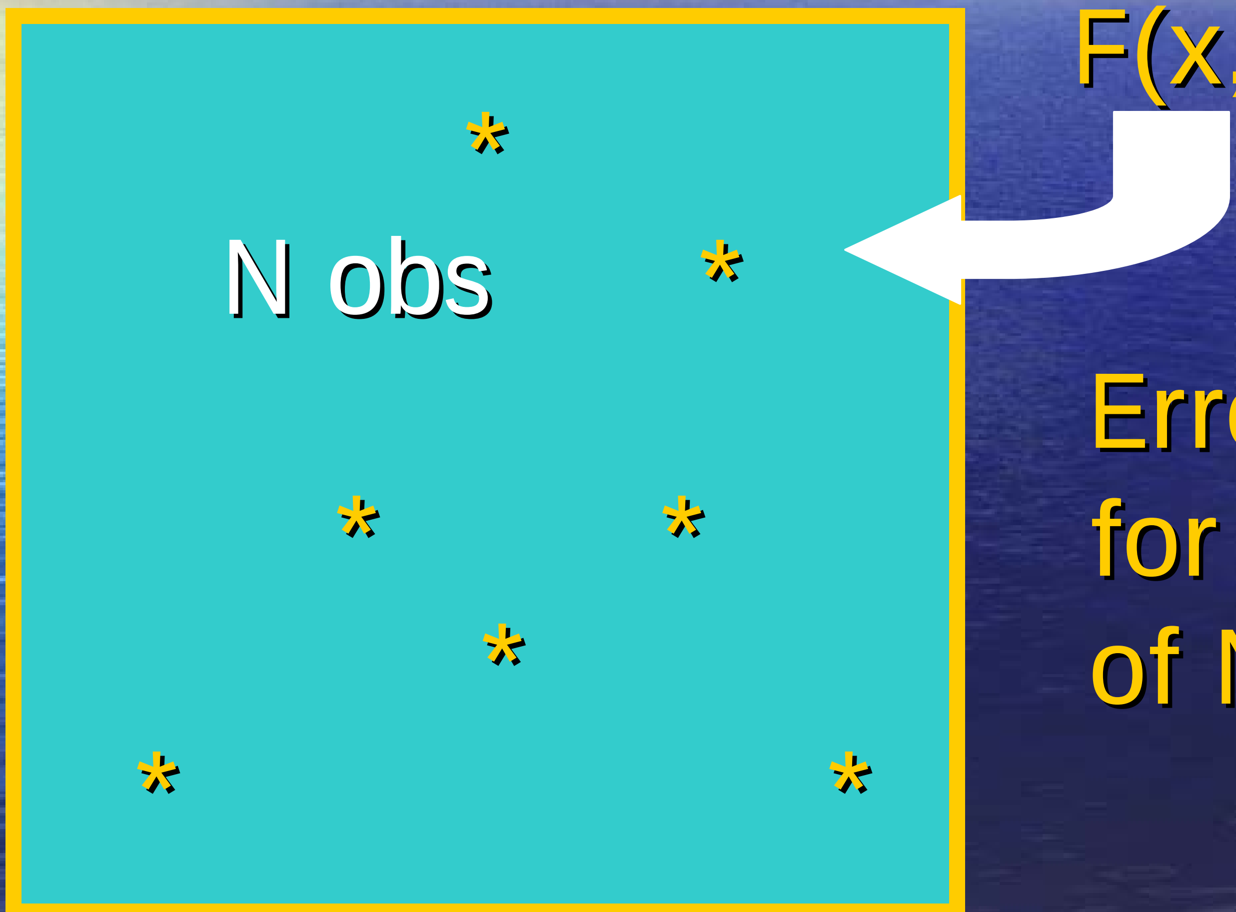
For sparsely sampled historical in situ data, a sparse grid (e.g. 5^o grid size) makes sense. Sparse grids best represent largest scales of variability. Hence the Reduced Space (RS) approach:

T = T' + T'' => C = C' + C'',

where T' is a linear combination of a few large scale patterns and C' is of low rank. If T' and T'' are assumed independent, there is a cheap way to compute T' and its error covariance P' corresponding to the OI solution cheaply (RSOI)

For a sparse grid observational error covariance R is usually assumed diagonal; its element are estimated as uncertainties in the grid box averages

F(x,y) [or F(x,y,t)]



Error variance for the mean of N observ is

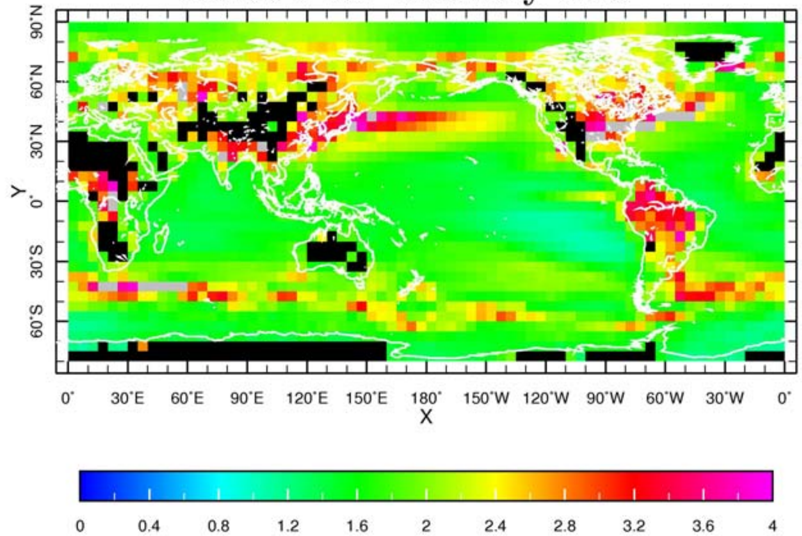
$$\sigma^2/N$$

Finer data coverage provided by satellite data is helpful for better estimates of R.

(right) modeling sampling error in grid box averages of in situ data

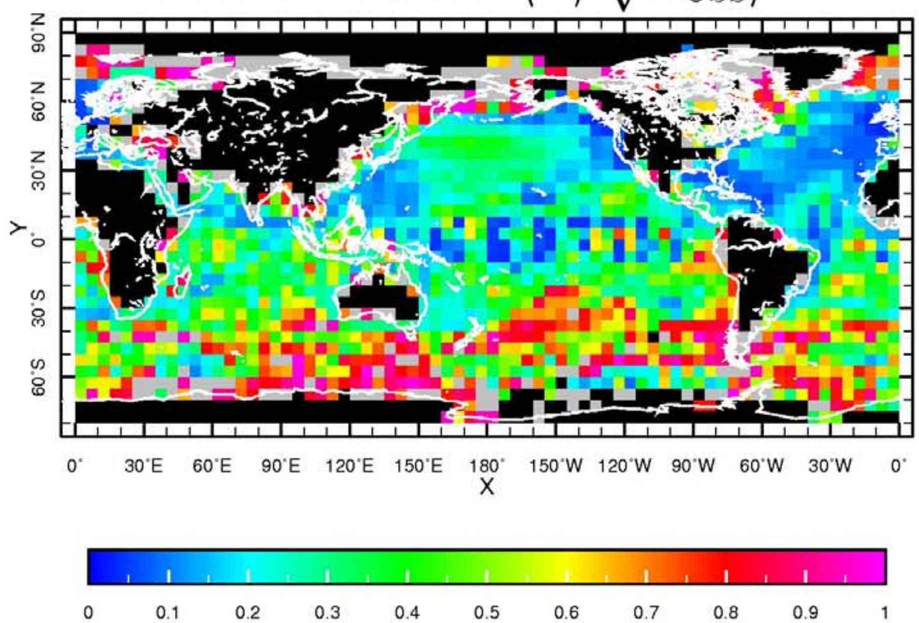
(bottom) verification

Single observation SST sampling+measurement error, ^{o}C, inside $5^{o}\times5^{o}$ monthly bins

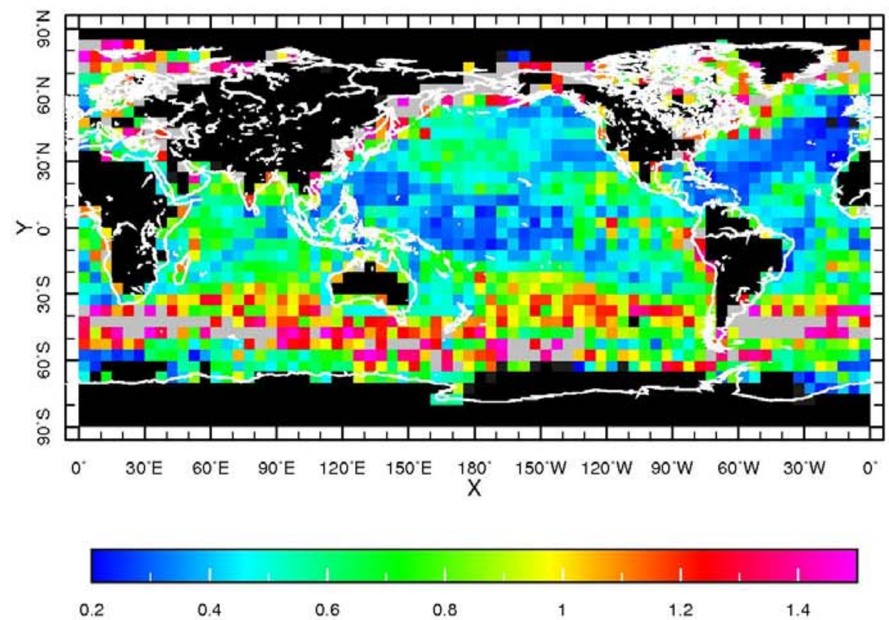


Modeling in situ data error for 5^{o} bins

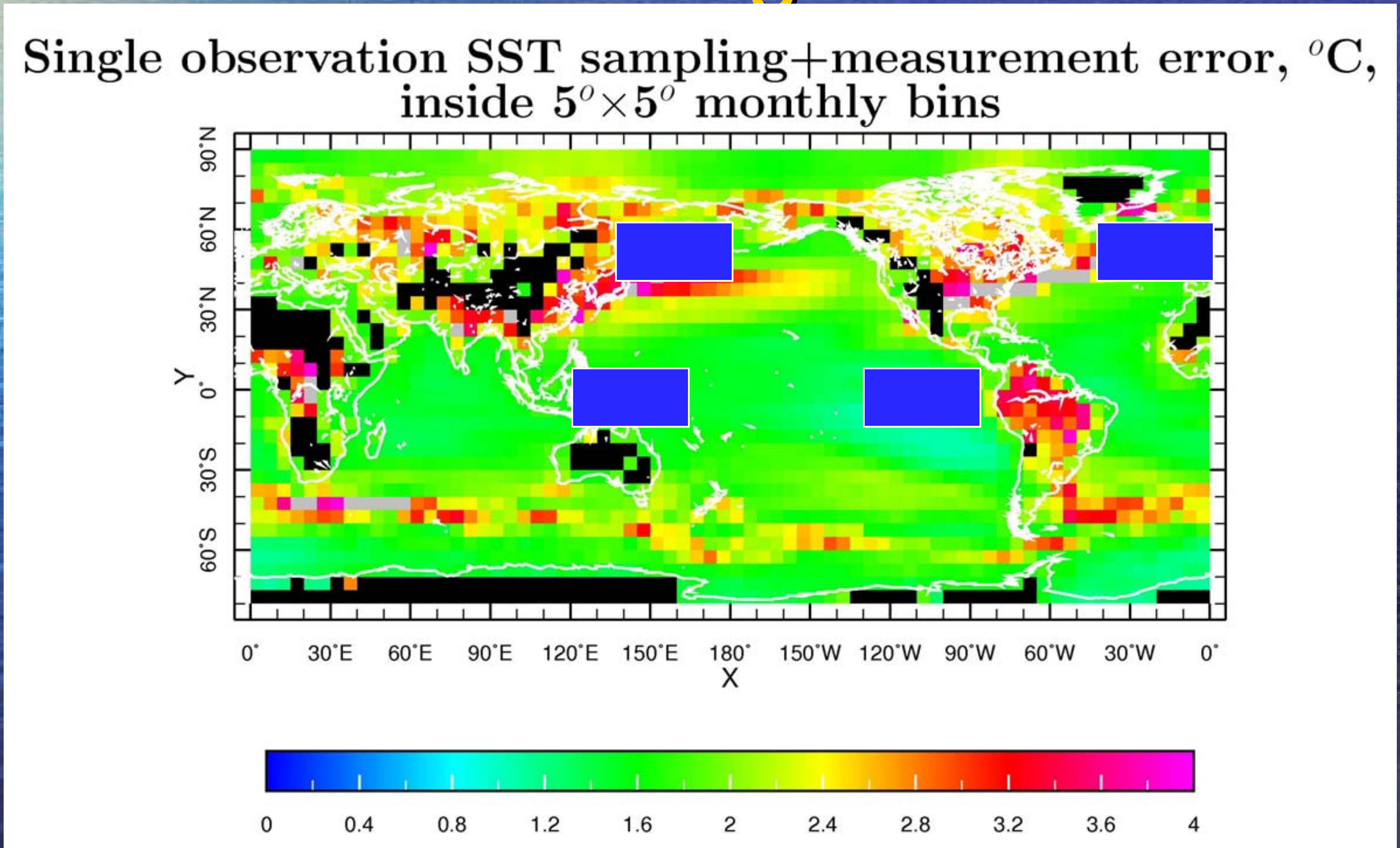
Modeled as $\langle \sigma / \sqrt{n_{\text{obs}}} \rangle$



Actual MODIS–ICOADS STD

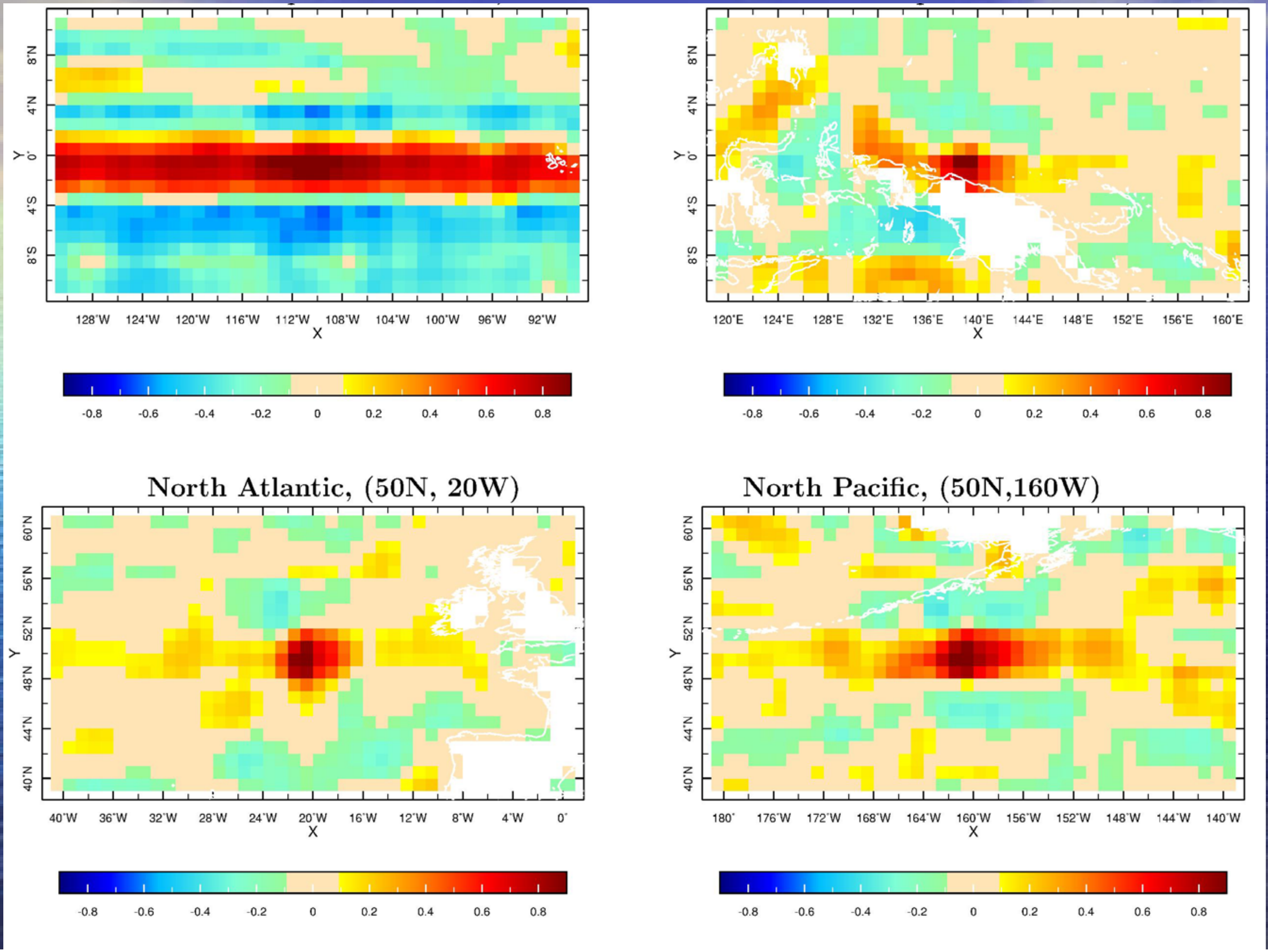


Finer data coverage is helpful for better estimates of R. If we would like to increase the resolution, a lot of detail is needed for estimating R and C”

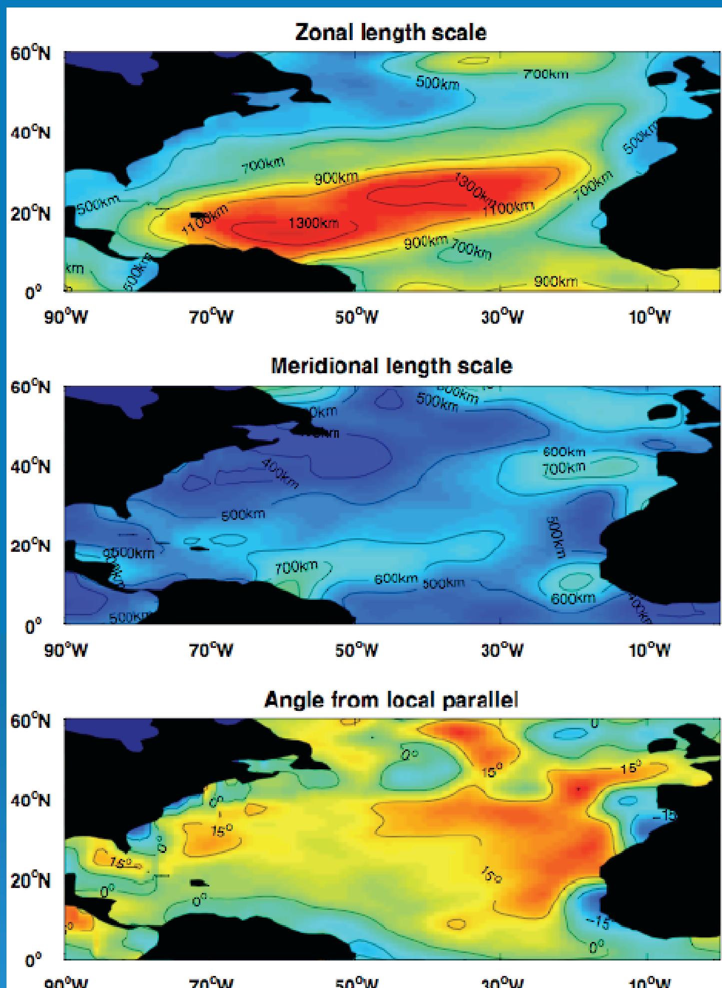


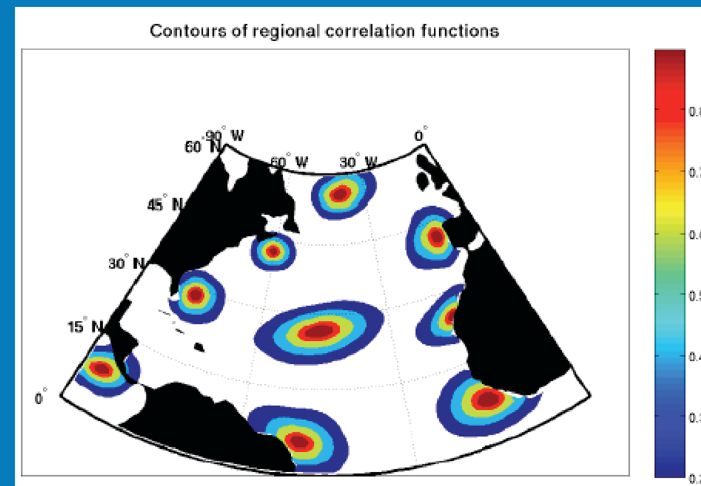
Truncation error autocovariance patterns:

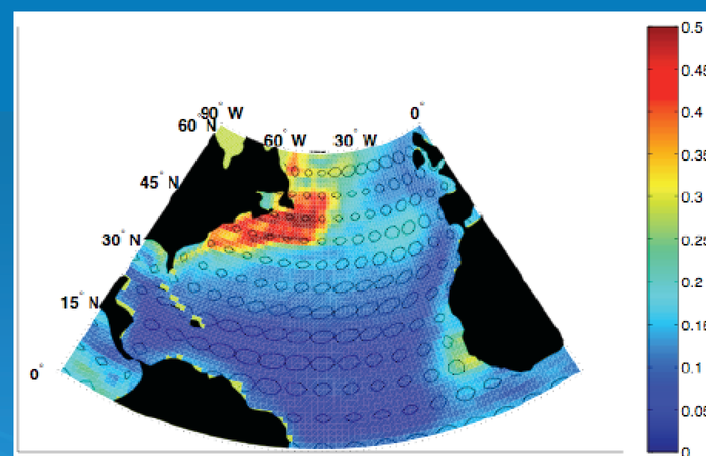
Eastern Equatorial Pacific at 110W   Western Equatorial Pacific at 140E



Covariance structures in the mid-scale SST

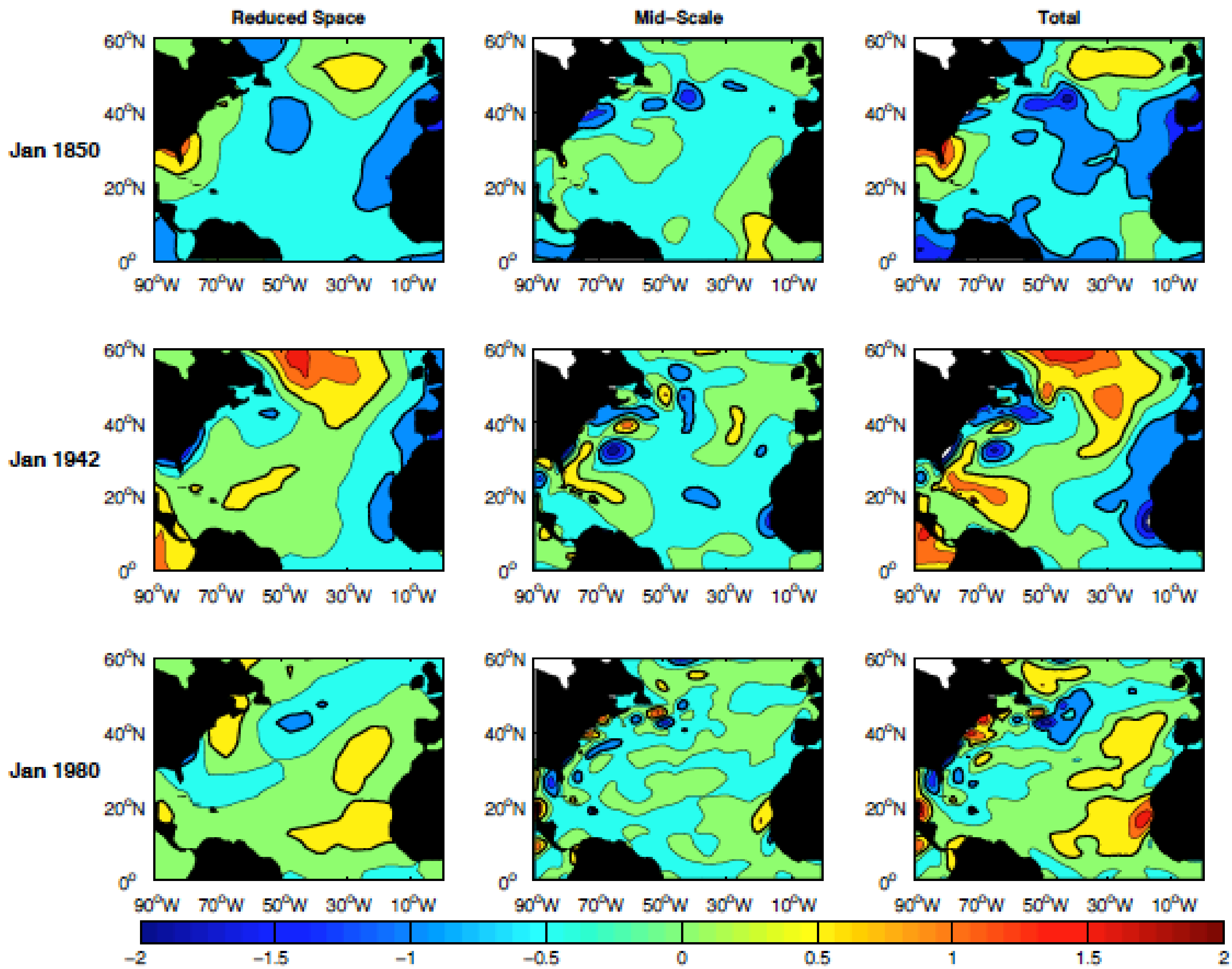




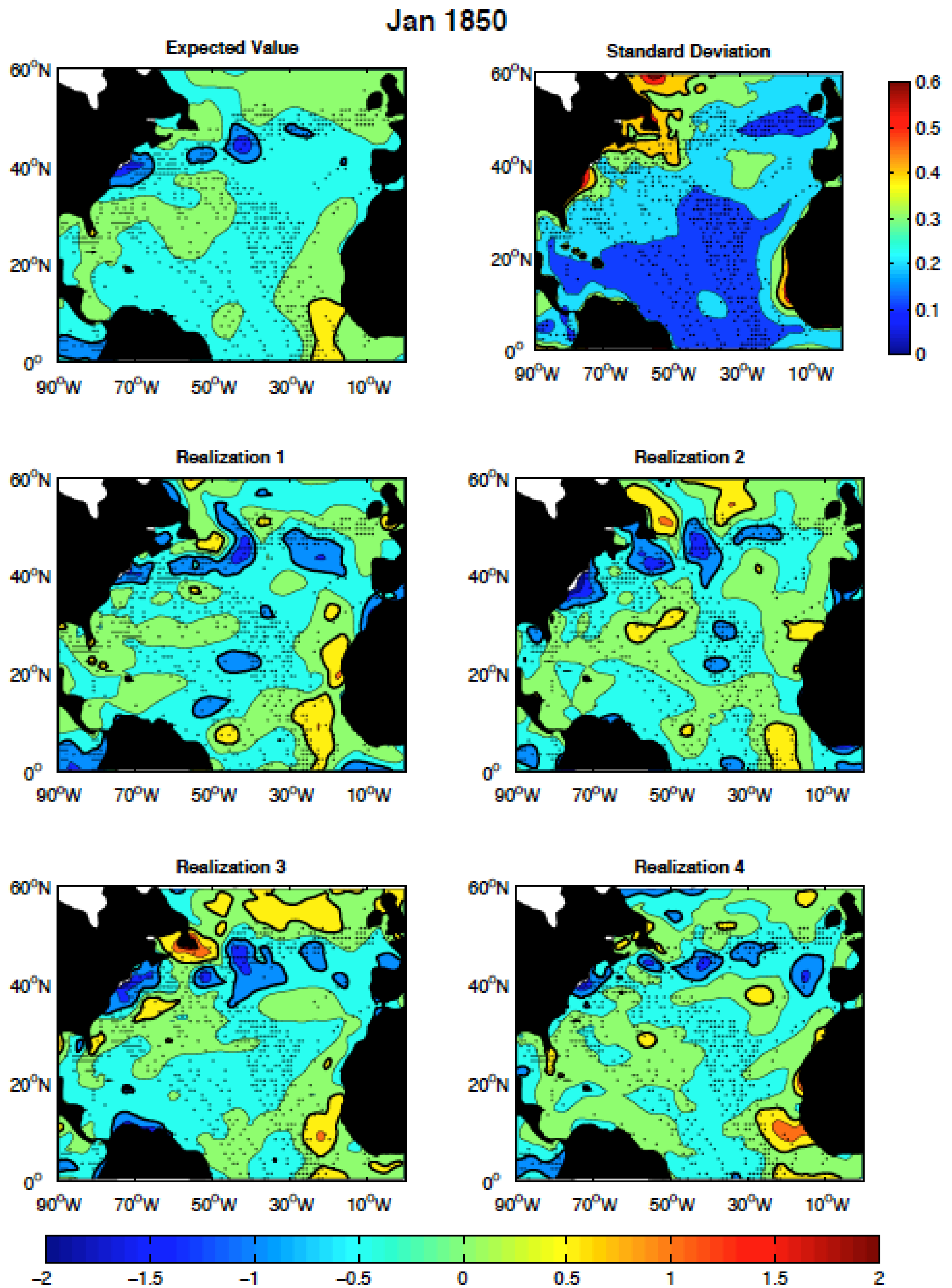


Parameters estimated from NCEP OI from 1981-present
via maximum likelihood

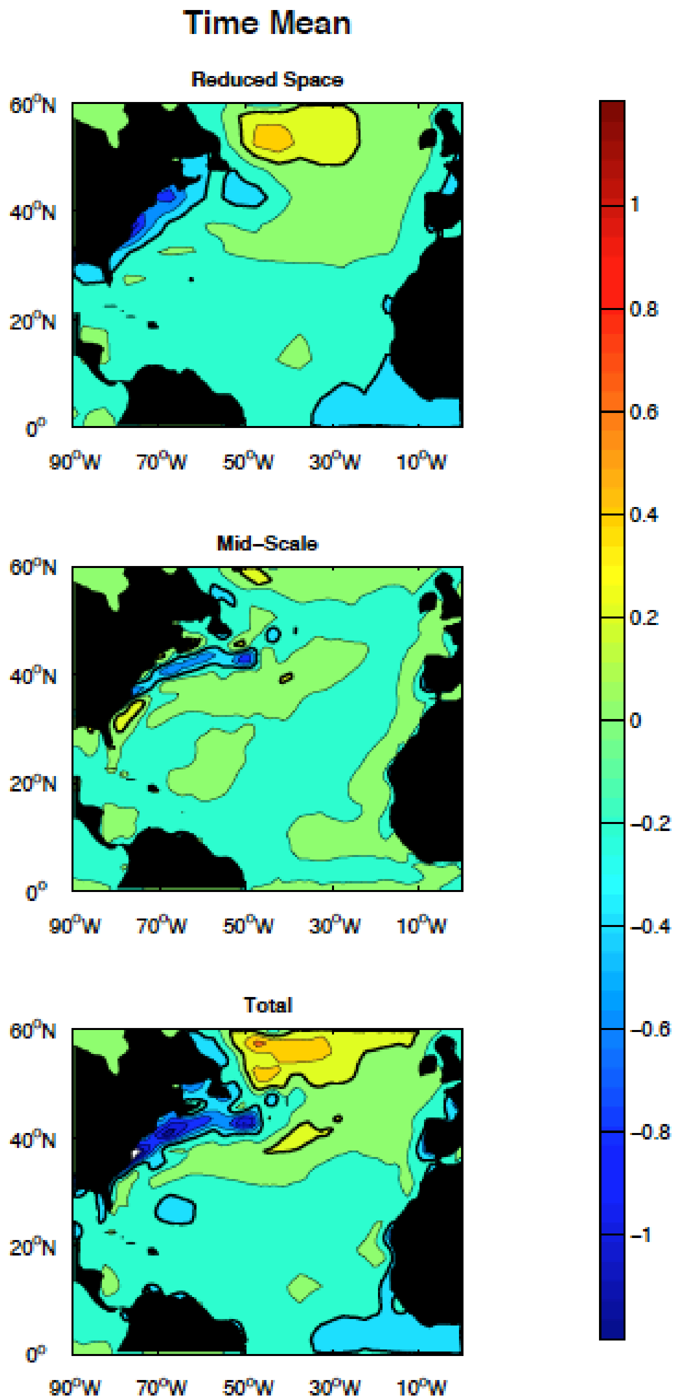
Karspeck et al. [2011, QJRMS, revised]

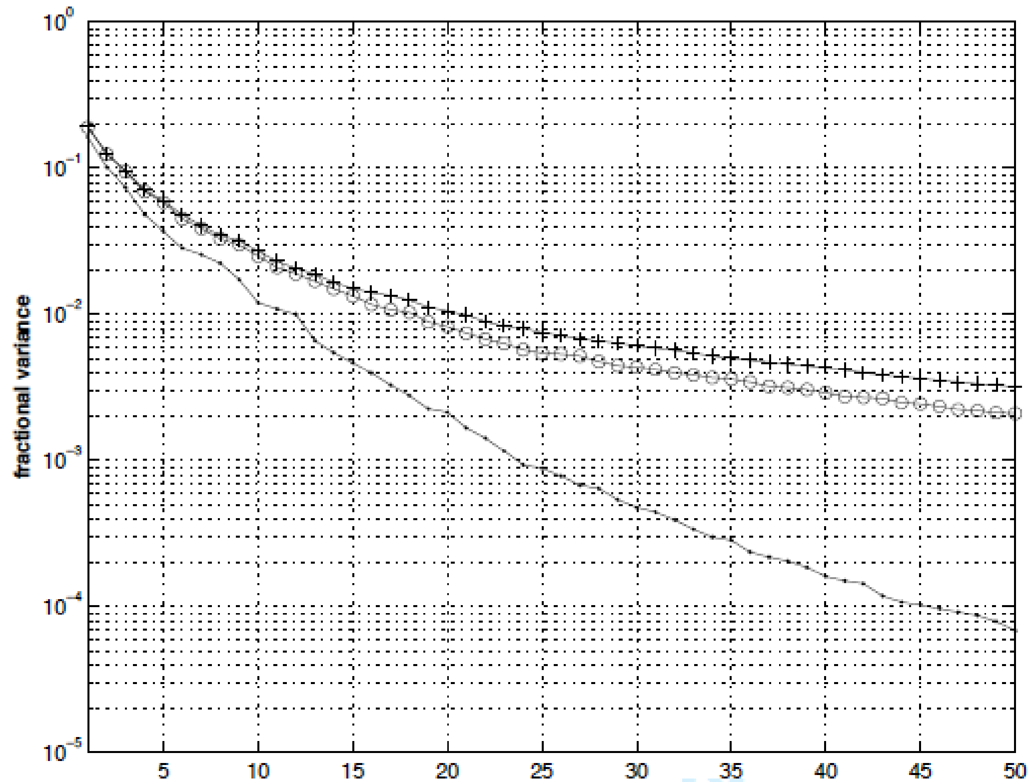


Karspeck et al. [2011, QJRMS, revised]



Karspeck et al. [2011, QJRMS, revised]



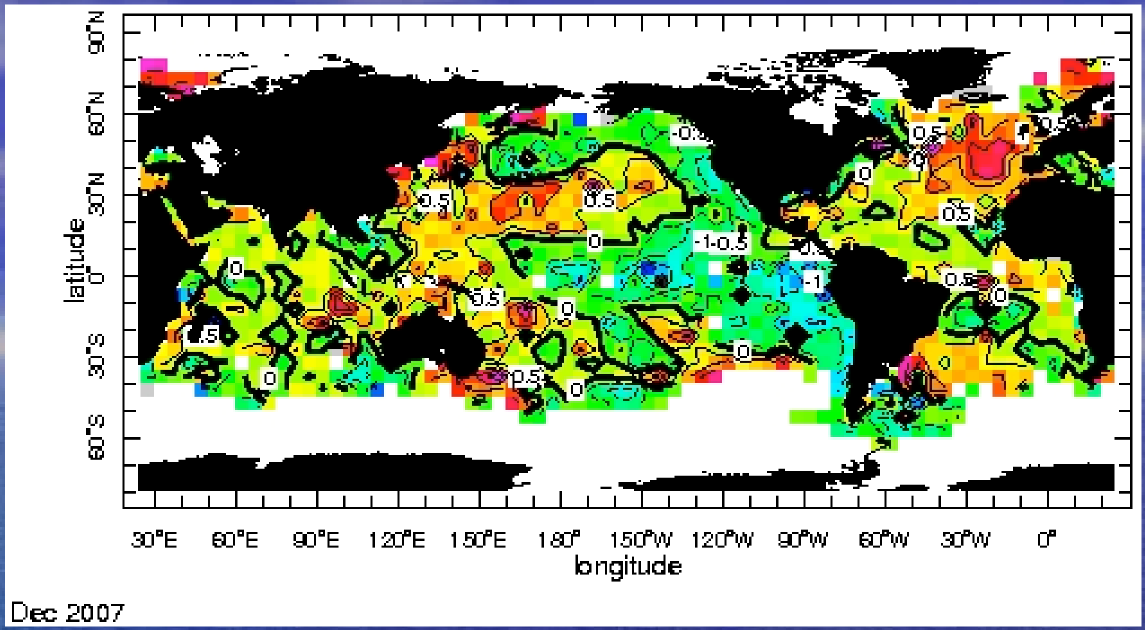


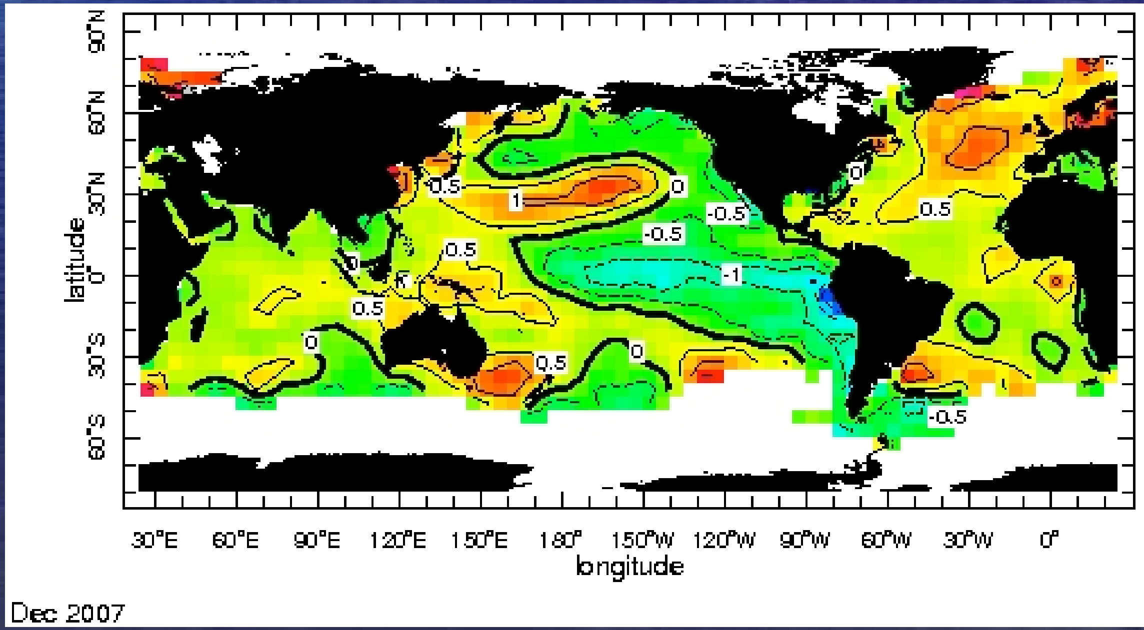
Karspeck et al. [2011, QJRMS, revised]

$$LT \sim N(LT_{OI}, LP_{OI}L^T)$$

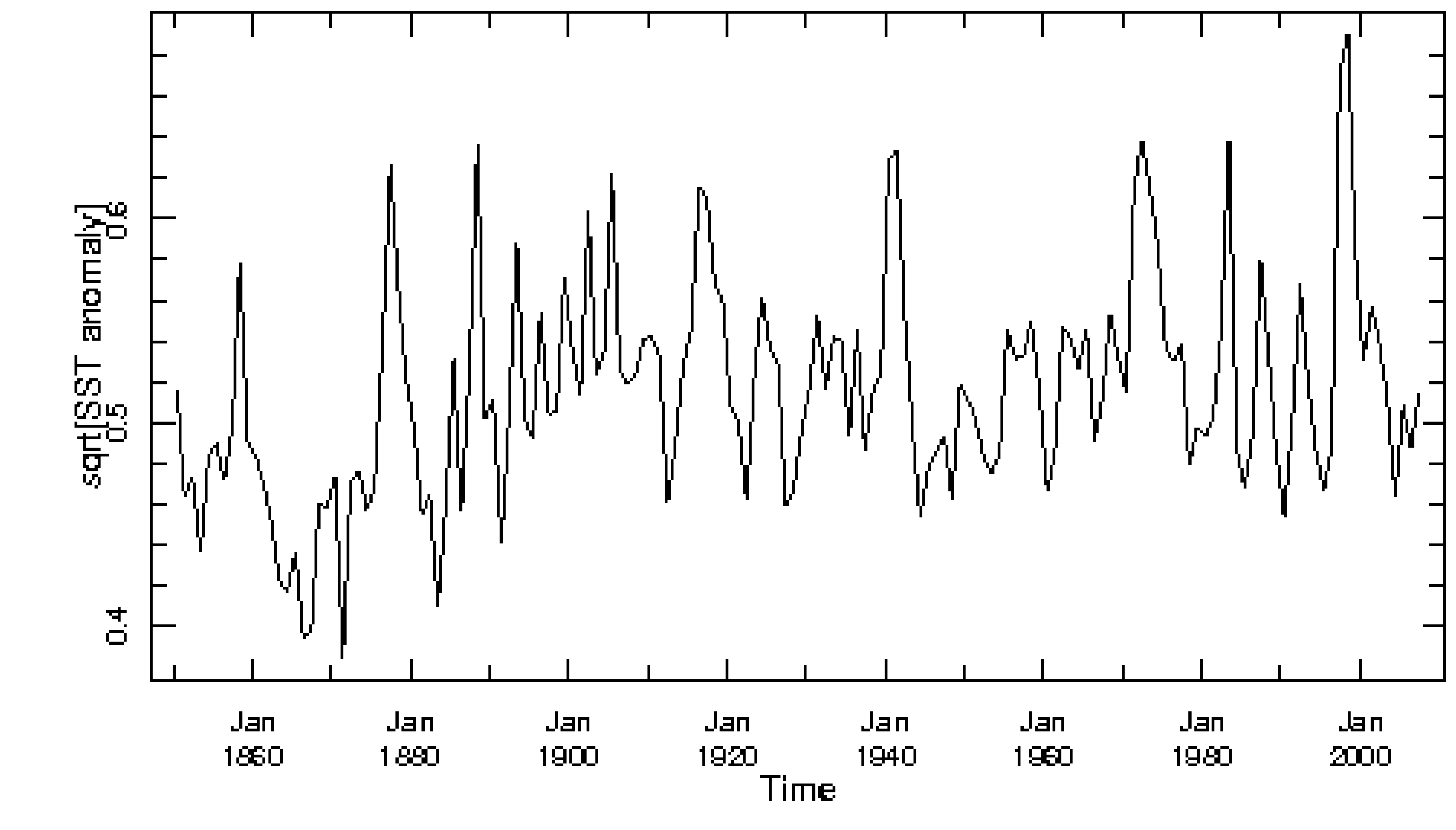
$$\langle T^T S T \rangle = T_{OI}^T S T_{OI} + Tr(SP_{OI})$$

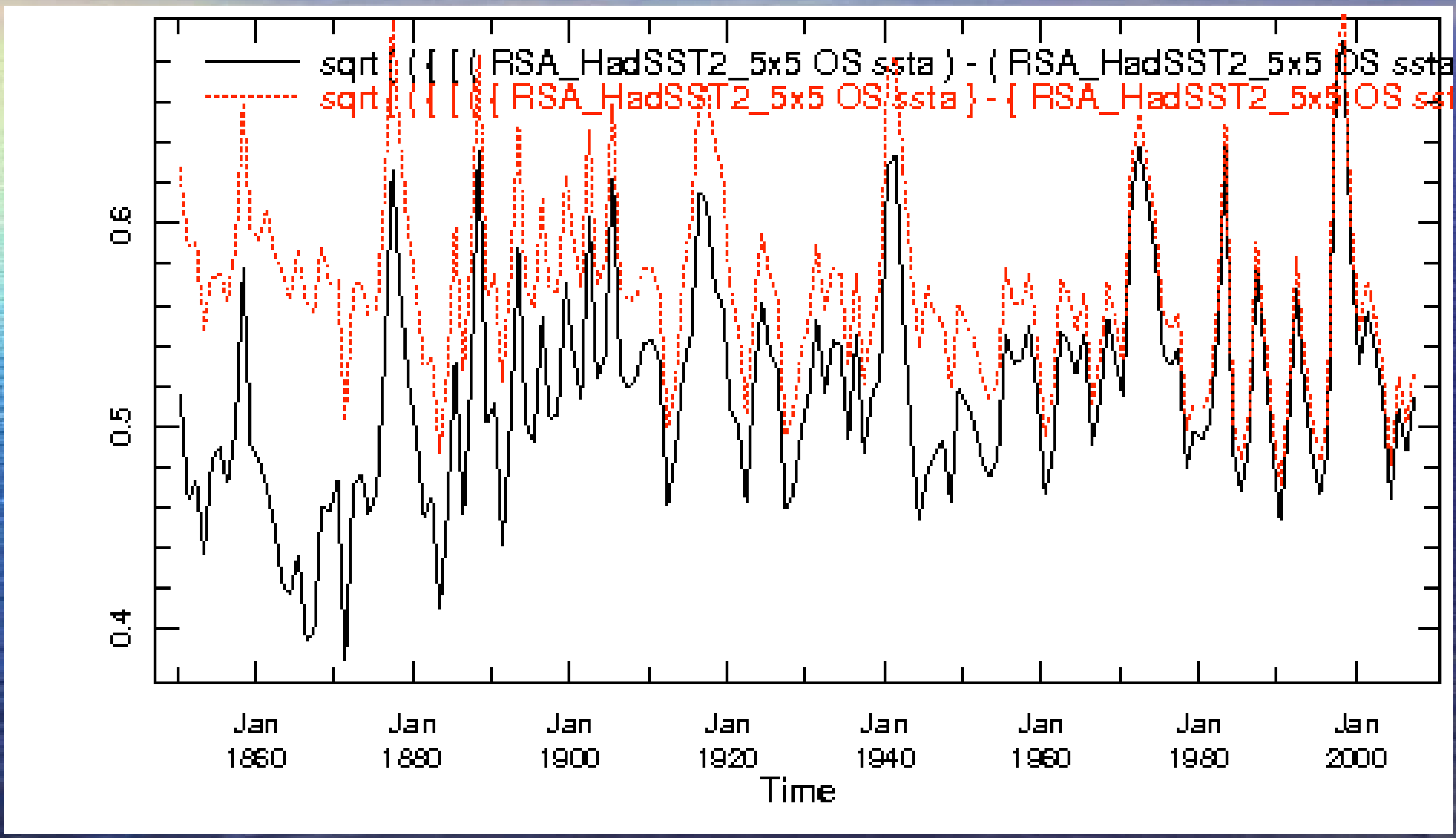
Reduced Space Optimal Smoother applied to HadSST2 data set of the U.K. MetOffice





Historical spatial variability of SST anomalies




sqrt (([(RSA_HadSST2_5x5 OS ssta) - (RSA_HadSST2_5x5 OS ssta
sqrt (([({ RSA_HadSST2_5x5 OS ssta } - { RSA_HadSST2_5x5 OS ss
0.6
0.5
0.4
Jan 1860
Jan 1880
Jan 1900
Jan 1920
Jan 1940
Jan 1960
Jan 1980
Jan 2000
Time

Conclusions:

1. High-resolution data satellite data can be used for estimating data covariance at different scales and then be used in efficient multi-scale optimal analysis procedures.
2. Data covariance is spatially non-stationary at both large and small scales and can be modeled as such.
3. Representation of the uncertainty by a set of samples from the posterior distribution is viable and useful

Supported by a NASA Physical Oceanography grant and an NSF CMG grant.