Extended Appendix to the paper

# Reduced Space Optimal Interpolation of Historical Marine Sea Level Pressure: 1854–1992

Alexey Kaplan, Mark A. Cane, and Yochanan Kushnir

Lamont–Doherty Earth Observatory of Columbia University

Palisades, NY 10964

Appendix is unpublished; Paper is in preparation

August 1998

## Appendix A: Re-estimation of the signal covariance

*Kaplan et al.* [1997] (hereafter K97) in their Appendix B obtained equations

$$A_p \stackrel{\text{def}}{=} \langle \alpha^p \alpha^{p\,T} \rangle = \Lambda + \mathcal{P}^p, \tag{A1}$$

$$A_{\text{OI}} \stackrel{\text{def}}{=} \langle \alpha^{\text{OI}} \alpha^{\text{OI}\,T} \rangle = \Lambda(\Lambda + \mathcal{P}^p)^{-1}\Lambda. \tag{A2}$$

which tie together covariances of the projection and reduced space OI solutions ($\alpha^p$ and $\alpha^{\text{OI}}$ respectively), error covariance for the projection solution $\mathcal{P}^p$, and the covariance of the retained portion of the signal space, $\Lambda$. The projection solution $\alpha^p$ consists of the best fit coefficients for the predetermined set of patterns (a basis of the reduced space, the columns of matrix $E$) to the observed data. $\mathcal{P}^p$ is the theoretical covariance of the error in these coefficients.

The covariance $\Lambda$ is the principal assumption in our computational procedure: it determines the distribution of energy over the basis of the reduced space. If the underlying assumptions of the method hold, the values of analyses covariances $A_p$ and $A_{\text{OI}}$ obtained from the solution should be approximately equal to the theoretical values given by the respective right-hand sides of the equations (A1, A2). In particular, the covariance of the projection solution should exceed the assumed covariance of the "true" signal, while the latter should exceed the covariance of the OI. In K97 we tested this consistency by plotting the diagonals of the matrices involved, particularly, through plotting the ratios of the diagonals of the analyses' covariances to the diagonal of $\Lambda$. We expect the projection ratio to be larger than unity and the OI ratio to be smaller than unity. If no adjustment is made to the covariance found through our standard procedure, this consistency check fails for the present SLP analysis: both ratios are less than unity (Figure A1(a)). This test did not hold for the unadjusted signal covariance in our SST analyses either, but a heuristic procedure was introduced which changed only the distribution of energy over the EOF modes and preserved the diagonality of $\Lambda$. This approach implicitly assumed that the individual EOFs modes could be used as they are, without any rotation. Results of this procedure were quite satisfactory for the SST analysis (see Figure 15 of K97), but failed for the present SLP analysis (Figure A1(b)). Consequently, here we develop a more complete and less heuristic procedure based on equations (A1, A2).

Inserting (A1) into (A2) to eliminate $\mathcal{P}^p$ yields

$$A_p = \Lambda A_{\mathrm{OI}}^{-1} \Lambda \tag{A3}$$

in which $A_p$ and $A_{\mathrm{OI}}$ are the known estimates from the projection and OI solutions, respectively, and $\Lambda$ is being sought in the class of symmetric non-negative definite matrices. An exact solution to the non-linear matrix equation (A3) can be constructed in the following way. We define canonical decompositions of the symmetric non-negative matrices $A_p$ and $A_{\mathrm{OI}}$ by

$$A_p = E_p S_p^2 E_p^T \ , \qquad A_{\mathrm{OI}} = E_{\mathrm{OI}} S_{\mathrm{OI}}^{-2} E_{\mathrm{OI}}^T \ , \tag{A4}$$

where $E_p$ and $E_{\mathrm{OI}}$ are orthogonal matrices, while $S_p^2$ and $S_{\mathrm{OI}}^2$ are diagonal matrices with non-negative elements. Inserting equations (A4) into (A3) we obtain

$$E_p S_p^2 E_p^T = \Lambda E_{\mathrm{OI}} S_{\mathrm{OI}}^2 E_{\mathrm{OI}}^T \Lambda \ ,$$

or

$$(E_p S_p)(E_p S_p)^T = (\Lambda E_{\mathrm{OI}} S_{\mathrm{OI}})(\Lambda E_{\mathrm{OI}} S_{\mathrm{OI}})^T$$

The latter can be true if and only if there exists an orthogonal matrix $U$ such that

$$E_p S_p U = \Lambda E_{\mathrm{OI}} S_{\mathrm{OI}},$$

and thus

$$\Lambda = E_p S_p U S_{\mathrm{OI}}^{-1} E_{\mathrm{OI}}^T. \tag{A5}$$

Now our task is to choose $U$ so that $\Lambda$ defined by formula (A5) is symmetric and non-negative definite. Symmetry means that

$$E_p S_p U S_{\mathrm{OI}}^{-1} E_{\mathrm{OI}}^T = E_{\mathrm{OI}} S_{\mathrm{OI}}^{-1} U^T S_p E_p^T,$$

or

$$U S_{\mathrm{OI}}^{-1} E_{\mathrm{OI}}^T E_p S_p^{-1} = S_p^{-1} E_P^T E_{\mathrm{OI}} S_{\mathrm{OI}}^{-1} U^T = (U S_{\mathrm{OI}}^{-1} E_{\mathrm{OI}}^T E_p S_p^{-1})^T. \tag{A6}$$

We define the singular value decomposition of $S_{\mathrm{OI}}^{-1} E_{\mathrm{OI}}^T E_p S_p^{-1}$ by

$$S_{\mathrm{OI}}^{-1} E_{\mathrm{OI}}^T E_p S_p^{-1} = G_1 \Sigma G_2^T \ ; \tag{A7}$$

$G_1$ and $G_2$ being orthgonal matrices, and $\Sigma$ being a diagonal matrix with non-negative elements. Obviously, the choice

$$U = G_2 G_1^T$$

makes

$$U S_{\mathrm{OI}}^{-1} E_{\mathrm{OI}}^T E_p S_p^{-1} = G_2 \Sigma G_2^T$$

symmetric and thus satisfies equation (A6). As a result we have

$$\Lambda = E_p S_p G_2 \Sigma G_2^T S_p E_p^T, \tag{A8}$$

which is obviously symmetric and non-negative (as $\Sigma$ is non-negative).

The entire procedure of the analysis then is performed in the following two-stage way. The covariance of the field is estimated, EOF patterns and eigenvalues are computed and used in projection and OI analyses without any adjustment. From the results of the analysis, $A_p$ and $A_{\mathrm{OI}}$ are estimated, as is the signal covariance $\Lambda$, computed according to formula (A8). Canonical decomposition,

$$\Lambda = F \Lambda_1 F^T,$$

shows that the EOF patterns found originally should be rotated by the operator $F$ and the distribution of energy over these rotated (but still orthogonal!) patterns should be given by the elements of $\Lambda_1$. EOFs not included into the reduced space are not being rotated, but their eigenvalues are modified according to the formula

$$\lambda_i^1 = \gamma \lambda_i + c, \qquad i = L+1, \cdots, M$$

where $\gamma$ and $c$ are constants which we defined from the two conditions: conservation of the total variance $v$ of the spectrum:

$$v = Tr[\Lambda_1] + \gamma(v - Tr[\Lambda_1]) + c(M - L),$$

and "no-jump" after the last reduced space eigenvalue:

$$\gamma \lambda_{L+1} + c = \lambda_L^1.$$

These conditions give the values of constants

$$\gamma = \frac{v - Tr[\Lambda_1] - \lambda_{L+1}(M-L)}{v - Tr[\Lambda_1] - \lambda_L^1(M-L)}, \qquad c = \lambda_L^1 - \gamma\lambda_{L+1}.$$

With this corrected estimate of covariance we rerun the analysis and check if equations (A1, A2) hold for $A_p$ and $A_{\mathrm{OI}}$ estimated from its results (i.e. if the results of the analysis are consistent with our a priori assumption on the signal variance). If not, we could do another iteration. However for this particular SLP applications the agreement was quite satisfactory. Figure A1(c) demonstrates it for the diagonal parts of the matrices.

To close the story of the covariance reestimation for the present SLP analysis, we would like to find out the technical reason why the heuristic algorithm of K97 failed for the present application (or, in other words, why the present algorithm was needed for the SLP analysis, but not for the SST analysis of *Kaplan et al.* [1998]). For this we run our new covariance reestimation algorithm for the SST analysis of *Kaplan et al.* [1998]. The explanation is given by Figure A2: the reduced space EOF rotation matrix $F$ for the SLP analysis has more off-diagonal structure than that for the SST analysis. Also the eigenvalues estimated by the present algorithm are much closer to those estimated by the K97 method for SST than for SLP. Figure A3 shows the same by presenting absolute values of the diagonal elements of rotation matrices $F$ in the top plot.

As discussed in K97, a realistic estimate of the signal covariance is of crucial importance for obtaining realistic theoretical error estimates (for example, if we assume too little energy in the truncated modes, the analysis becomes overoptimistic about its ability to constrain the solution within the reduced space, and gives anything but conservative error estimates for the solution). When we repeat the experiment with the North Atlantic withheld area as in K97 and *Kaplan et al.* [1998] we find that (Figure A4) the theoretical error estimate in the middle of the North Atlantic withheld area reaches 0.8 mb, while when all the available data are used, the estimated error is about 0.3 mb. Consistent with these error estimates, the rms difference between the solutions with and without the data in the chosen area almost reaches 0.6 mb. This example is evidence that the analysis with the given settings produce reliable conservative error estimates for the solution.

## References

Kaplan, A., Y. Kushnir, M.A. Cane, and M.B. Blumenthal, 1997: Reduced Space Optimal Analysis for Historical Datasets: 136 years of Atlantic Sea Surface Temperatures. *J. Geoph. Res.*, **102**, 27,835–27,860.

Kaplan, A., M.A. Cane, Y. Kushnir, A.C. Clement, M.B. Blumenthal, B. Rajagopalan, 1998:. Analyses of global sea surface temperature 1856-1991. J. Geophys. Res., **103**, 18,567–18,589.

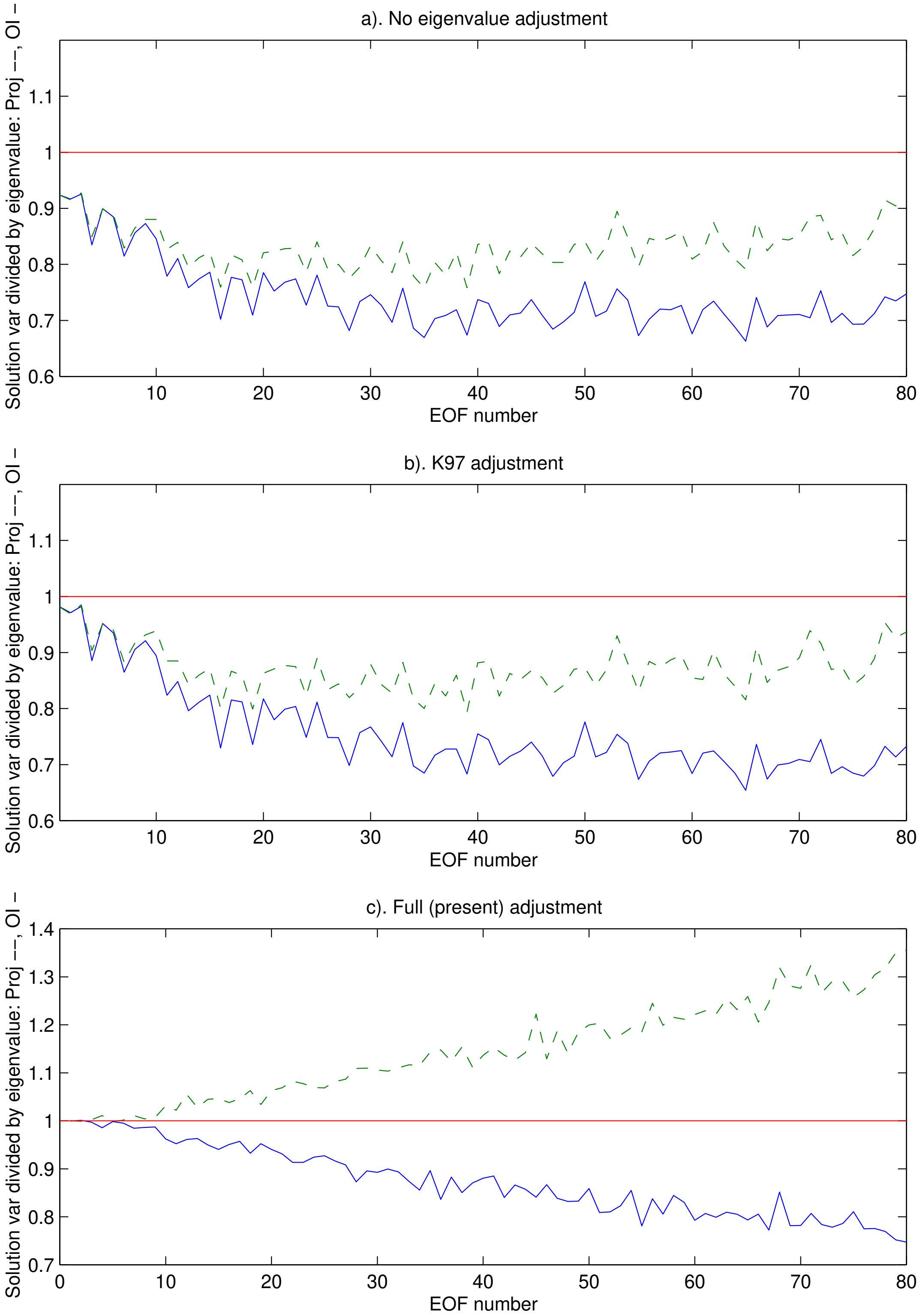
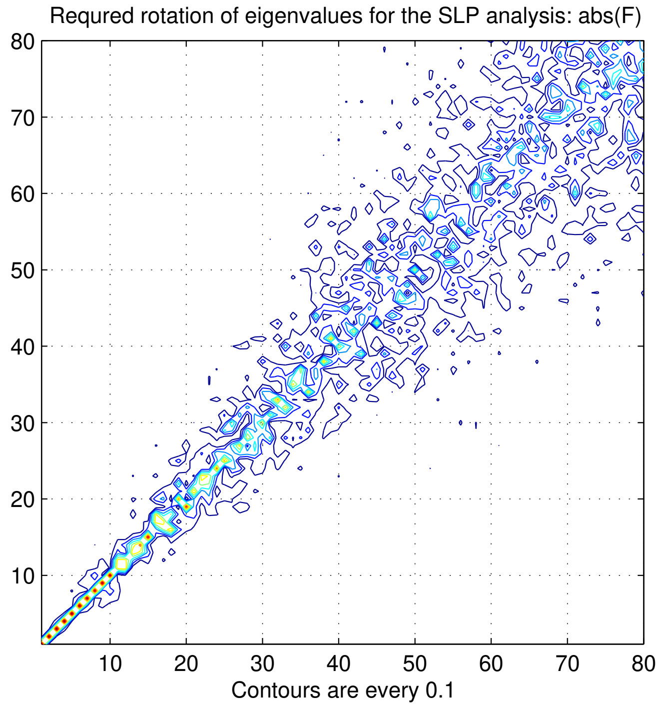



**Figure A1.** Performance of different adjustment schemes for the SLP analyses.



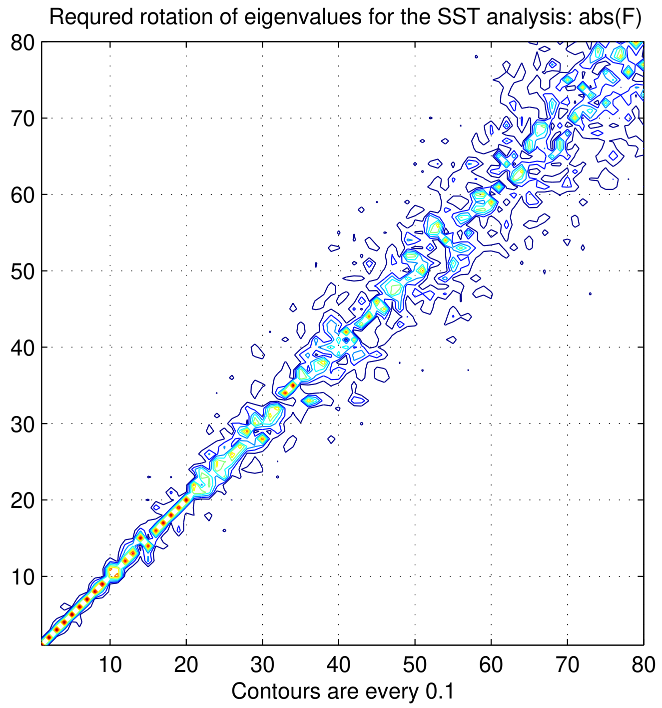




**Figure A2.** Absolute values of rotation matrices $F$ for SLP and SST analyses. Unities on diagonals and zeroes elsewhere would mean no rotation.

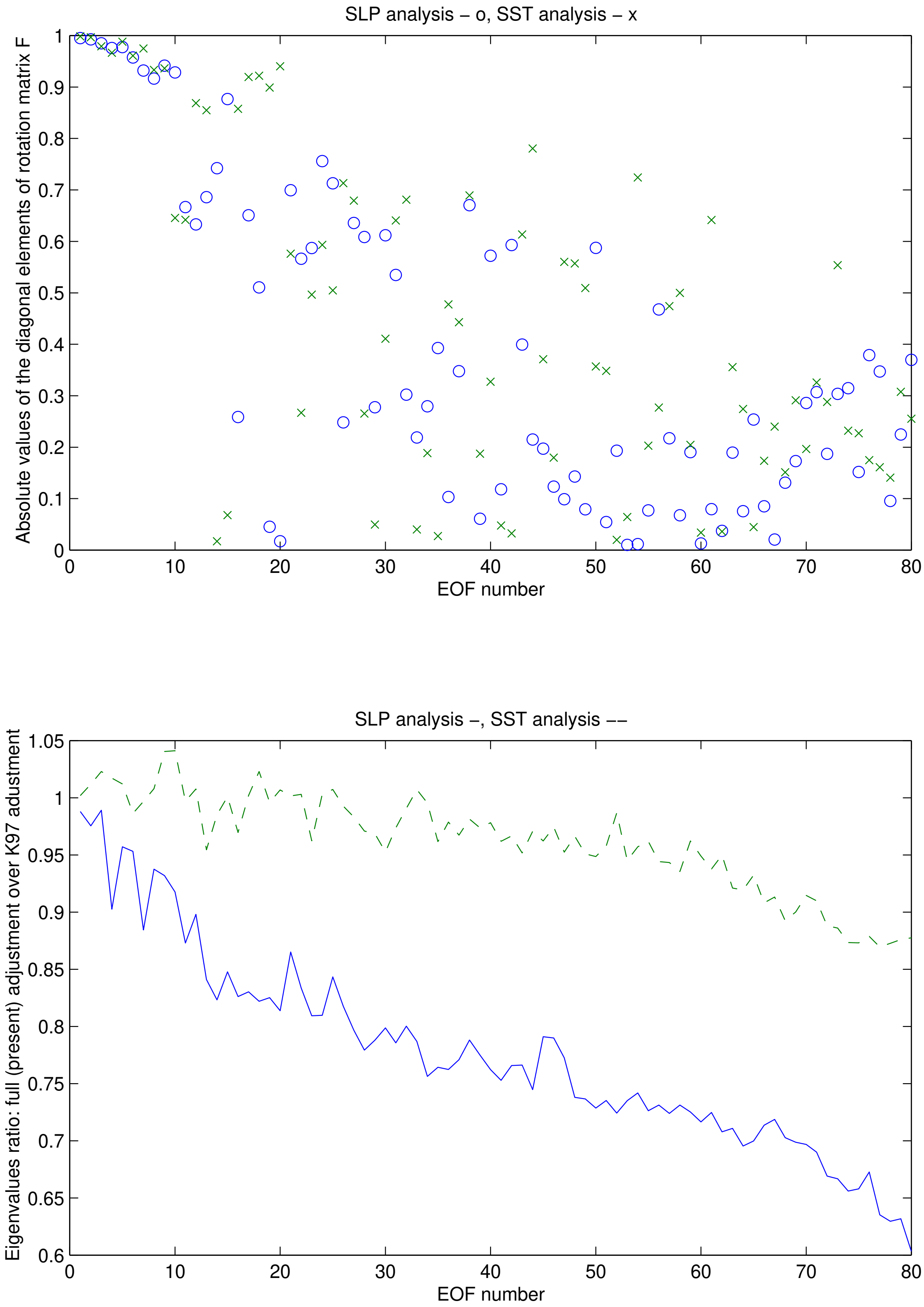
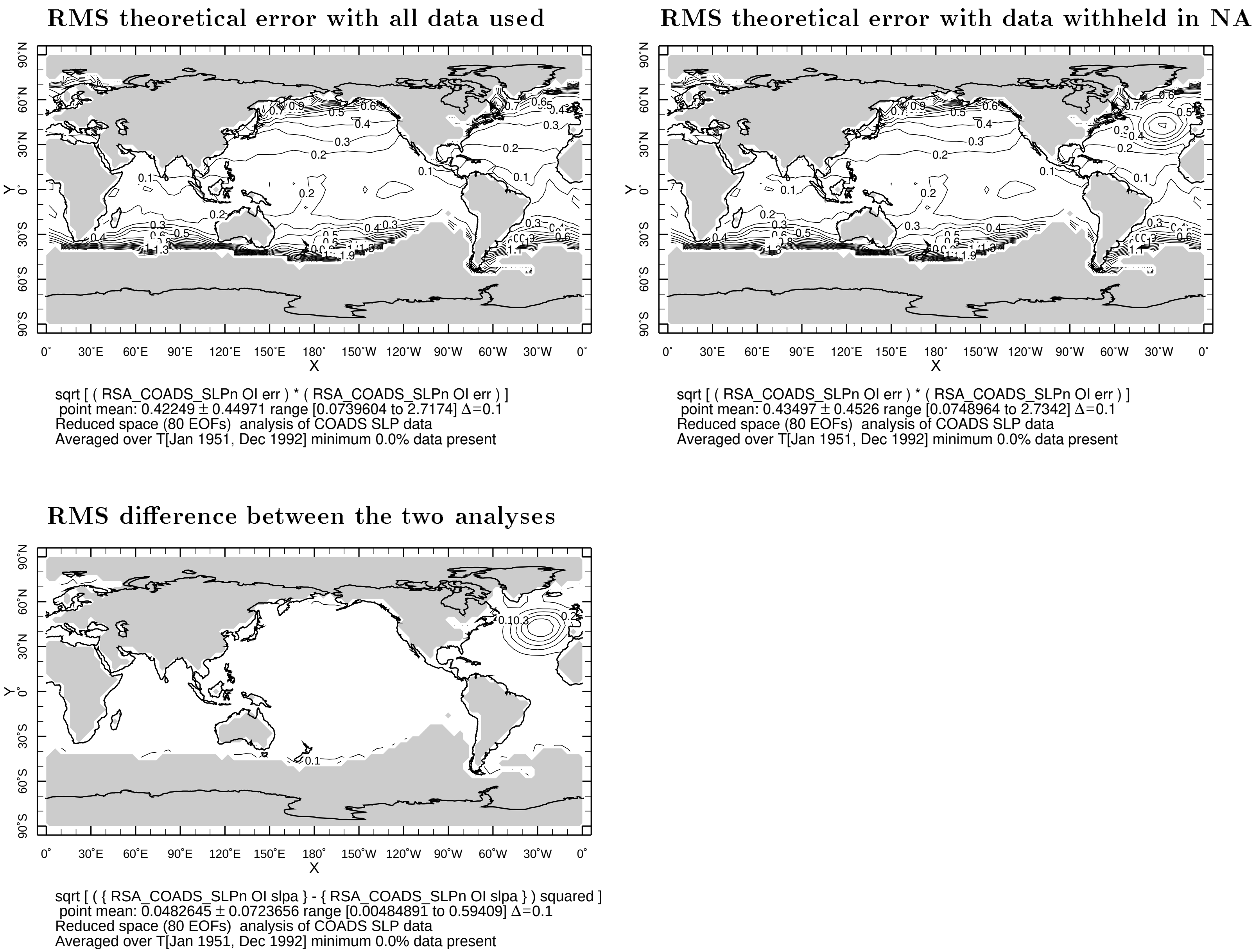



**Figure A3.** Comparison of the diagonals of rotation matrix $F$ (top) and estimated eigenvalues (bottom)




FigureA4 Experiment with the withheld North Atlantic area and its comparison with the standard analysis run.