MULTI-TAPER RESPONSE ESTIMATION FOR SATELLITE INDUCTION STUDIES

Catherine G. Constable

Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, University of California at San Diego, La Jolla, Ca92093-0225, USA. Email: cconstable@ucsd.edu

ABSTRACT

We assess the possibilities of obtaining improved low frequency electromagnetic response functions for Earth's mantle by using robust multi-taper spectral estimation to analyze long time series of satellite magnetic field observations. These responses are used to determine the 1-dimensional electrical conductivity profile of the mantle, and the longest period estimates provide constraints for the deepest part of the mantle. A test data set is used to show that outliers in the time series can have a deleterious impact on response estimates. Bias introduced as a result of inadequate knowledge about the relative noise distributions between internal and external field contributions is most significant at short periods and elsewhere when the coherence is low. When used on the same test data set multi-taper response estimates produce essentially the same results as section averaged spectral estimates, but improvements are still needed in estimating uncertainties for the multi-taper response functions.

1. INTRODUCTION

Satellite magnetometer data from Magsat, Ørsted, CHAMP and SAC-C have been used to determine the one-dimensional electromagnetic response in the Earth to variations in large scale magnetospheric sources ([1], [2], [3], [4], [5]). The electromagnetic responses are estimated in the form of a frequency dependent transfer function $Q(f)$ between the external magnetospheric ring current variations and the corresponding induced part. The usual representation has been

$$I(f) = Q(f)E(f) \tag{1}$$

or the equivalent in the time-domain,

$$i(t) = q(t) * e(t). \tag{2}$$

The external source field of interest is usually taken to be P_1^0 in geomagnetic coordinates aligned with Earth's dipole axis and, along with its internal counterpart, is isolated either by a global time-varying internal/ external field separation (often involving additional terms) or on a pass by pass basis. This provides the sampled time series of $i(t)$ and $e(t)$. A variety of cross spectral methods have been used to estimate the complex and frequency dependent $Q(f)$, which is generally transformed to Weidelt's [6] C-response prior to interpretation in terms of electrical conductivity. The conversion for a 1D, spherically symmetric conductivity distribution responding to a single spherical harmonic variation field of degree l is given by

$$c_l(f) = a\frac{l-(l+1)Q_l}{l(l+1)(1+Q_l)} = \frac{a(1-2Q_1(f))}{2(1+Q_1(f))} \tag{3}$$

2. A TEST DATA SET

We use a test data set kindly supplied by Nils Olsen, from an early analysis of 2001 CHAMP, Ørsted and Ørsted-2 (SAC-C) data described in [4], who modeled $e_1^0(t)$ and $i_1^0(t)$, in a global model with the temporal variation parametrized in cubic B-splines with knots at 4 hour intervals. The time series for e_1^0 and i_1^0 are shown in Fig. 1. Although these data series have since been substantially improved and extended [7], they are adequate to illustrate the methodology described here.

To calculate the C-response, Olsen et al. subdivide the record into sections, derive L independent estimates for the auto S_{ii}, S_{ee} and cross spectra S_{ie}, S_{ei} between internal and external contributions and then compute $Q(f)$ using averages of the L sections, before transforming to $C(f)$. If the external field time series $e(t)$ were assumed to be noise free and all the noise attributed to the internal estimate, then the equivalent least squares estimate of the transfer function would be just

$$Q(f) = \frac{< S_{ie} >}{< S_{ee} >}. \tag{4}$$

However both $e(t)$ and $i(t)$ are contaminated with noise, and in this case Eq. (4) leads to a biased estimate for $Q(f)$. A total least squares approach is preferable. When the relative noise fraction in each power spectrum is known, Olsen [8] discussed how to correct this bias. If

$$\zeta_i = \frac{\sigma_i^2}{S_{ii}} \qquad \text{and} \qquad \zeta_e = \frac{\sigma_e^2}{S_{ee}} \tag{5}$$

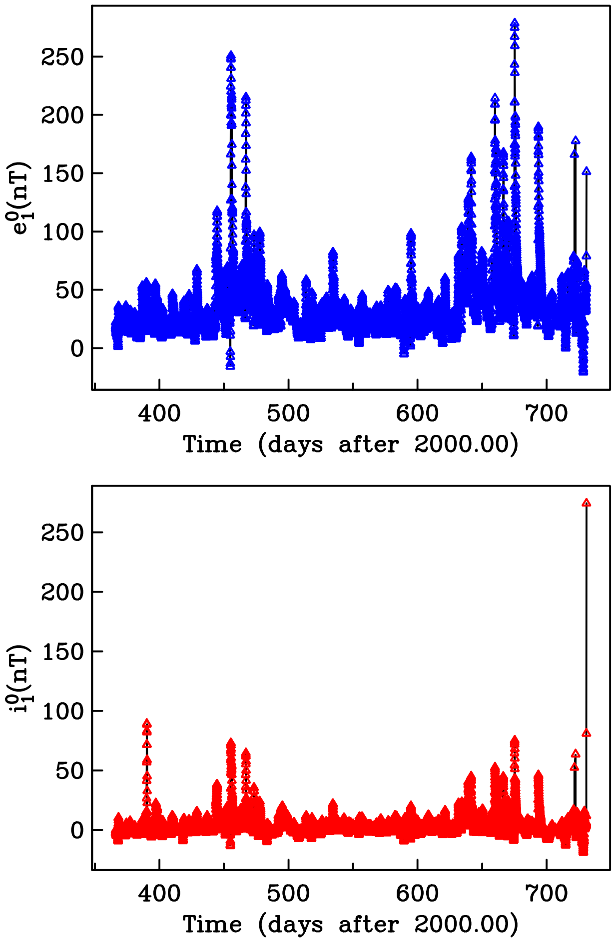

Figure 1. Time series for e_1^0 and i_1^0 used as a test data set

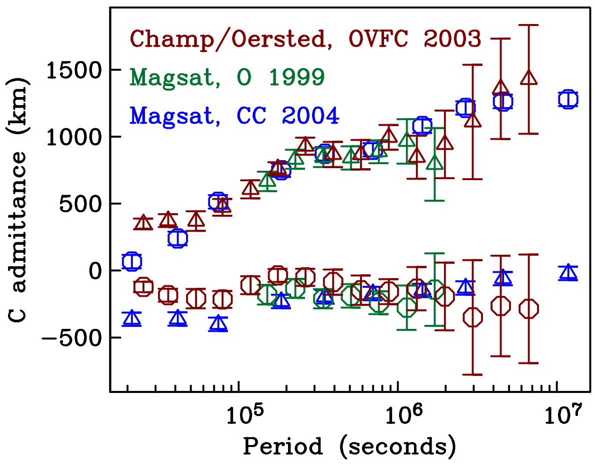

Figure 2. C-response estimates for the test data set described in Section 2 (dark red symbols are from [4]), results from Magsat using the same method (green, [1]), and multitaper estimates from Magsat data (blue, [2]).

then the relative noise in the two series can be written as

$$\eta = \frac{\zeta_e}{\zeta_i}, \tag{6}$$

and the estimate given by Eq. (4) can be corrected for the influence of noise in $e(t)$, yielding

$$Q(f,\eta) = \frac{Q(f)}{2}\left\{1 - \eta + \sqrt{(1-\eta^2) + \frac{4\eta}{\gamma^2}}\right\} \tag{7}$$

with γ^2 the frequency dependent coherence between $e(t)$ and $i(t)$

$$\gamma^2(f) = \frac{< S_{ie} >< S_{ei} >}{< S_{ee} >< S_{ii} >} \tag{8}$$

When $\eta = 1$, corresponding to equal relative noise in the external field and induced response, Eq. (7) reduces to the geometric average of the response given in Eq. (4) with the equivalent response obtained with the roles of the internal and external field time series reversed. That is

$$Q(f,1) = \sqrt{\frac{< S_{ie} >< S_{ii} >}{< S_{ee} >< S_{ei} >}} \tag{9}$$

as used by [4]. Q(f,1) is then transformed to $C(f)$ and shown as the dark red symbols in Fig. 2.

3. WHY USE MULTI-TAPER SPECTRAL ANALYSIS?

A remaining challenge in 1-dimensional induction studies is to extend response estimates to the longest periods possible, as these provide information about electrical conductivity structure of the deep mantle. The longest periods attainable are set by the length, T, of the time sections used, which in the section averaging approach is determined by the number L of independently analyzed sections. The frequency resolution Δf is proportional to $\frac{1}{T}$. Multitaper spectral analysis uses orthogonal tapers to sample the whole time series; each tapered sample provides an independent spectral estimate, these are then averaged to reduce the variance of the spectral estimates. The frequency resolution depends on the number and spectral properties of the tapers used. Averaging estimates from many tapers will give a smooth spectrum with lower frequency resolution (and the possibility of spectral leakage or bias), few tapers yield a high resolution estimate at the cost of greater uncertainty.

Two kinds of multitaper spectral estimation are in common use: one based on minimizing broadband bias advocated by Thomson [9] uses the prolate spheroidal family of tapers; the second [10] has been used in an approximate data adaptive fashion to minimize local loss (bias squared plus variance). The tapers for the second approach are closely approximated by sine functions which are easy to calculate.

3.1. Adaptive Estimation and Frequency Resolution

Use of a locally adaptive method means that both the resolution and the error can vary in the frequency domain. When k sinusoidal multi-tapers are used in the estimate the frequency resolution for the estimate $\hat{S}(f)$ is

$$\Delta f = k f_N / N_f \quad (10)$$

with f_N the Nyquist frequency, and N_f the number of frequencies estimated. The adaptive part of the procedure is the optimal choice of k. Riedel & Siderenko (1995) give an asymptotic result for minimizing the local loss (bias squared plus variance) of

$$k_{opt} \sim \left[\frac{12 S(f) N^2}{S''(f)} \right]^{2/5}. \quad (11)$$

S'' can in principle be determined from a quadratic fit to a pilot estimate of the local power spectral density, but the above expression must be modified to restrict the rate of growth of the number of tapers near regions where S'' vanishes. The spectra from the individual sine tapers are averaged, with parabolic weighting that tapers to zero beyond the kth spectral estimate. This weighting scheme ensures a smooth spectrum as the number of tapers changes with frequency (Robert Parker, personal communication, 2006). In practice this has proved hard to implement for cross spectral analysis, the resulting responses have poor frequency resolution and we have instead used a fixed number of tapers. Constable & Constable [2] opted to use 20 tapers to analyze the Magsat data series, yielding a uniform frequency resolution of Δf = 20/(2188*7200)= 1.3 μHz. The blue and green symbols in Fig. 2 show a comparison of multi-taper responses to section averaged responses, respectively, for Magsat. There are some minor differences between response estimates from the section averaging and multi-taper methods at both long and short periods. One goal of this work is to elucidate the source of these differences by testing methods on exactly the same data series, something which has not previously been done. A second goal not yet achieved is to assess whether it is possible to reduce the rather large uncertainties in the section averaged estimates.

4. EXPERIMENTS WITH THE TEST DATA

We computed a multi-taper response estimate using the test data series of Fig. 1, using the same assumption about the distribution of noise between i and e as used in [4], namely $\eta = 1.0$. The estimates are plotted in blue in Fig. 3 and are in good agreement with the section averaging results except at the very shortest period. Other differences are that the multi-taper estimates extend to longer period, and have smaller nominal uncertainties. These should not be considered reliable, especially at long periods where they are based on standard deviations among averages across frequency bands containing individual estimates that must be correlated, based on our expected frequency resolution.

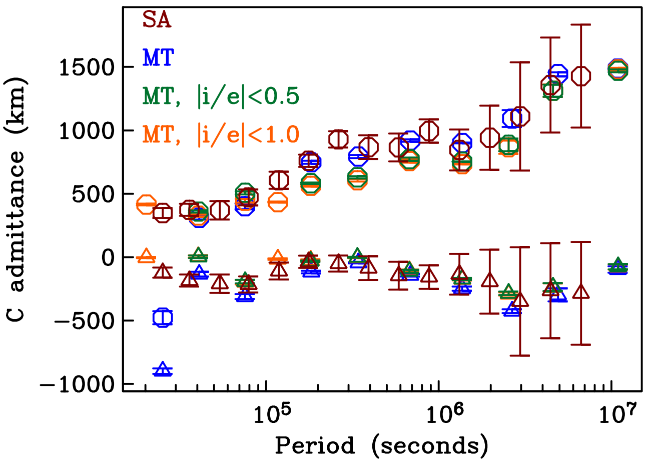

Figure 3. Comparison of section averaged response (SA, dark red) with multi-taper (MT, blue). Orange and green symbols show the effect of eliminating data for which $|\frac{i(t)}{e(t)}| > 1.0$ *and* $|\frac{i(t)}{e(t)}| > 0.5$*, respectively. In all cases* $\eta = 1.0$.

4.1. Influence of outliers in the time series

The average ratio for $|\frac{i(t)}{e(t)}|$ is about 0.3, although it varies widely along the time series as can be seen from Fig. 4. We tested the effect of eliminating all data pairs for which $|\frac{i(t)}{e(t)}| > 1.0$ on the grounds that the induced field should always be smaller than the external field. This eliminates slightly $< 10\%$ of the data and has a substantial influence on the response estimates which appear to have been biased upwards by the presence of this noise, especially in the $10^5 - 10^6$s period range. More drastic measures, eliminating all $|\frac{i(t)}{e(t)}| > 0.5$ have little further impact (Fig. 3) in this particular data set.

4.2. Assumptions about noise in $i(t)$ and $e(t)$.

We also tested the possible influence of bias due to incorrect assumptions about the relative noise levels in $i(t)$ and $e(t)$ (by varying η in Eq. (7). We computed response estimates under the assumption of (i) no noise in $e(t)$, corresponding to $\eta = 0.0$, (ii) equal relative noise in $e(t)$ and $i(t)$, $\eta = 1.0$, and (iii) $\eta = 0.25$, internal estimates have twice the relative standard error of external. As expected from Eq. (7) the effects are most pronounced in those frequency bands where the coherence is lowest. Effects are relatively small except at periods near one day and its harmonics where departures from assumptions about the source field are likely to be most significant (Fig. 5).

4.3. What to do about uncertainty estimates

Improved uncertainty estimates are needed for the multi-taper response estimates, especially at long periods. A

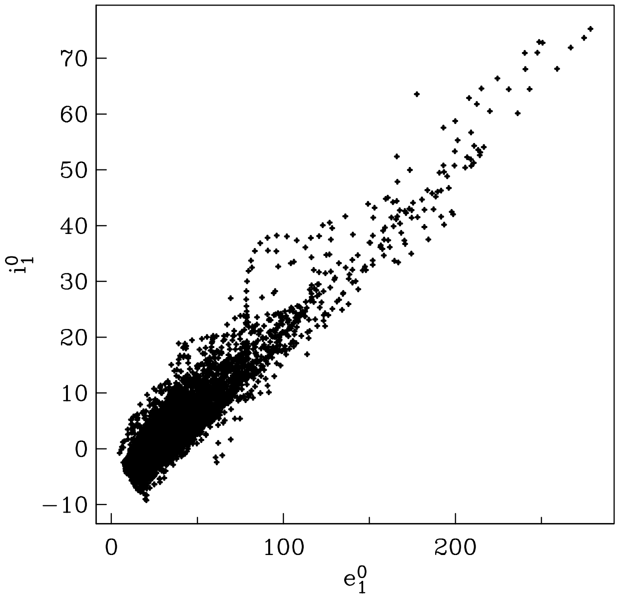

Figure 4. Scatter plot of internal i_1^0 versus external e_1^0 field contributions at the same time points for the test data set of Fig. 1.

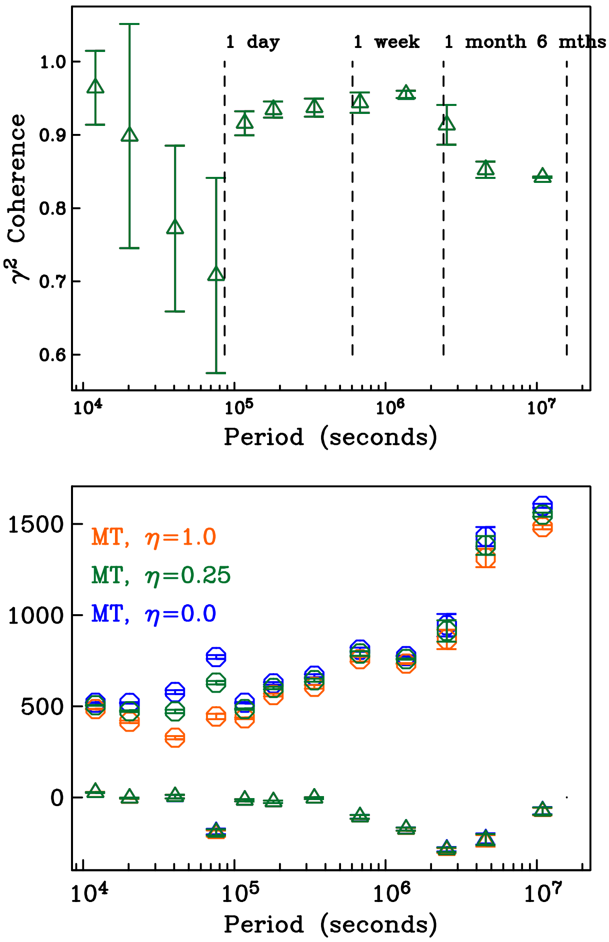

Figure 5. Comparison of responses computed under various assumptions about partitioning noise between i and e. In all cases data with $|\frac{i(t)}{e(t)}| > 1.0$ have been rejected.

jackknife approach based on the individual taper estimates is under investigation. As in section averaging method the individual spectral estimates can be regarded as independent, and can be expected to provide more realistic estimates than calculated to date.

5. SUMMARY

These experiments suggest that it is important to use robust processing to compensate for outliers in the time series and avoid bias in electromagnetic response estimates. The results do not appear very sensitive to knowing the exact relative distribution of noise between internal and external time series, unless the coherence is low. Electrical structure in the deep mantle is controlled by the long period variations in the response (especially the imaginary part), making it important to provide realistic error bars in this regime. This is the subject of ongoing efforts.

ACKNOWLEDGMENTS

Thanks to Nils Olsen for supplying the test data and section averaged response estimates from [4], and to Bob Parker who developed the sine multi-taper cross-spectral analysis software used here. This work was supported under NASA Grant NAG5-13747.

REFERENCES

[1] N. Olsen. *Surv. Geophys.*, 20:309–340, 1999.

[2] S. Constable and C. Constable. *Geochem. Geophys. Geosyst.*, 5:Q01006DOI 10.1029/2003GC000634, 2004.

[3] G. Balasis, G. Egbert, and S. Maus. *Geophys. Res. Lett.*, 31:L16610, doi:10.1029/2004GL020147, 2004.

[4] N. Olsen, S. Vennerstrøm, and E. Friis-Christensen. *First CHAMP Mission Results for Gravity, Magnetic and Atmospheric Studies*, C. Reigber and H. Lühr and P. Schwinzer (Springer, Berlin):245–250, 2003.

[5] N.C. Richmond, C.G. Constable, S. Constable, and J. Ribaudo. *1st Swarm International Science Meeting*, 1:116, 2006.

[6] P. Weidelt. *Zeit. Geophys.*, 38:257–289, 1972.

[7] A. Kuvshinov and N. Olsen, 2006.

[8] N. Olsen. *Geophys. J. Int.*, 133:298–308, 1998.

[9] D.J. Thomson. *IEEE Proc.*, 70:1055–1096, 1982.

[10] K. Riedel and A. Sidorenko. *IEEE Transactions on Signal Processing*, 43:188–195, 1995.