BIRRP: Bounded Influence, Remote Reference Processing

Revision History

Revision 1 Oct 2001 added robust ar filter computation as default

Revision 2 Nov 2001 added basic and advanced input options, input of all components from one ascii file,use of dimensional period instead of Nyquists for input and output, output of apparent resistivity and phase

Revision 3 Mar 2002 numerous bug fixes in output section, provision of both period and frequency in output, allowance of sample rate as well as sample interval, change to scale estimate for Thomson weights

Revision 3.1 fixed minor errors in rarfilt

Revision 3.2 changed convergence flag to integer

Revision 3.3 added immediate termination if all weights go to zero

Revision 4 Dec 2002 major improvements to robust ar filter routine, added option to input prewhitened data and prewhitening filter

Revision 5 Mar 2004 Major revisions throughout the code

Revision 5.1 minor bug fixes

Changed filter so that frequency is input in dimensional units instead of Nyquists, added output in j-format,integrated selection of delete-one files with diagnostics, added sort feature to output so that frequencies will be monotonically increasing, changed input when nfil<0 so that dimensional frequency, real and imaginary filter coefficients are input and fit with a taut cubic spline, added option to input data start and window start/end times (instead of points to skip and read) and let program automatically align time series, added output of qq plot and diagnostic file for first stage in two stage processing, removed l1 norm initial solution option, reduced memory usage by ~50%, added output of stage 1 transfer functions and coherences with jackknife error bounds WARNING--the order of parameters input to the program and some of the output filenames are different in Version 5 than in earlier versions!

This document describes the latest revision of the code BIRRP for bounded influence remote reference magnetotelluric data processing. This version contains several major enhancements to earlier robust mt processing codes. The program is provided for research purposes only on the understanding that it will not be used in any commercial application or included in any set of programs that are to be sold. Please report bugs to:

Dr. Alan D. Chave
Woods Hole Oceanographic Institution
Woods Hole, MA 02543
+508-289-2833 (voice)
+508-457-2006 (fax)
alan@whoi.edu

The BIRRP program computes magnetotelluric and geomagnetic depth sounding response

functions using a bounded influence, remote reference method, along with an implementation of the jackknife to get error estimates on the result. It also incorporates a method for controlling leverage points (magnetic field values which are statistically anomalous), the implementation of two stage processing which allows for removal of outliers in both the local electric and magnetic field variables, and allows multiple remote reference sites to be used. The bounded influence algorithm is described in:

Chave, A.D., and D.J. Thomson, A bounded influence regression estimator based on the statistics of the hat matrix, *J. Roy. Stat. Soc., Series C (Appl. Stat.)*, 52, 307-322, 2003.

The extensions required for magnetotelluric analysis are described in:

Chave, A.D., and D.J. Thomson, Bounded influence estimation of magnetotelluric response functions, *Geophys. J. Int.*, *157*, 988-1006, 2004.

Both of these papers may be requested via the author's website http://www.whoi.edu/science/AOPE/people/achave/. An earlier and somewhat outdated version of the procedure is described in:

A.D. Chave and D.J. Thomson, Some comments on magnetotelluric response function estimation, *J. Geophys. Res., 94*, 14215-14225, 1989.

For a tutorial background on robust spectral analysis, please see:

A.D. Chave, D.J. Thomson, and M.E. Ander, On the robust estimation of power spectra, coherences, and transfer functions, *J. Geophys. Res., 92*, 633-648, 1987.

The application of the jackknife to spectral analysis problems is covered in:

D.J. Thomson and A.D. Chave, Jackknife error estimates for spectra, coherences, and transfer functions, in S. Haykin (ed.), *Advances in Spectral Analysis and Array Processing*, Englewood Cliffs: Prentice-Hall, pp. 58-113, 1991.

In addition, a code to perform electric and magnetic field galvanic distortion tensor decomposition of the magnetotelluric response tensor output by BIRRP is available. Called GBDECOMP, this requires the delete-one response files output by BIRRP. Technical details may be found in

A.D. Chave and J.T. Smith, On electric and magnetic galvanic distortion tensor decompositions, *J. Geophys. Res., 99*, 4669-4682, 1994.

BIRRP is based on section-averaging of a long time series. It differs from some of my earlier robust codes in that no band averaging is performed because of difficulties this causes the jackknife due to correlation. Instead, the section length is variable and selected frequencies within each section are processed to yield the response. Experience has shown that bounded influence

algorithms are most sensitive to outliers when the section length is of order one over the frequency of interest, and hence superior performance is achieved using variable section lengths. The processing proceeds as follows: after reading in the data, they are prewhitened using an autoregressive filter (computed robustly, which is essential) and the Fourier transforms of the sections are computed after tapering with a prolate data window. Corrections for prewhitening are made and any instrument high or low pass filters may be applied at this point with a user-supplied external function FILTER described later. A series of rotations may be applied to first convert the sample coordinates to an orthogonal system, then rotate it to a new direction. The response functions and associated prediction coherences are then computed using both conventional and bounded influence remote reference methods. The user may choose either to use conventional remote reference processing or the new two stage bounded influence procedure. Estimates of the standard error on the result are computed using the unbalanced jackknife at a specified group of frequencies. The section length is then reduced and the process is repeated. A series of diagnostics may optionally be output. These include the data for quantile-quantile plots at selected frequencies for both the original and bounded influence regressions and data for plots of normalized input, output, and residual power and the hat matrix diagonal. The final response functions are stored in separate files for each component, and the prediction coherences are also in their own files. These files are named in a logical manner using a root supplied by the user. A set of large binary files containing the delete one estimates of the response functions is optionally output for use by GBDECOMP which performs galvanic distortion decompositions on the tensor.

BIRRP is written in ANSI standard Fortran 77 with the following exceptions:

1.Some variable names are longer than 6 characters.
2.The $ delimiter to suppress terminal carriage return has been used in format statements.

The code has successfully been ported from a Sun Sparcstation 2 running SunOS 4 to an UltrabookIIi running Solaris 7 to Hewlett Packard workstations and a SunBlade 1000 running Solaris 8 without modification.It also has been ported to MacOSX Absoft Fortran and to various Windows machines. The distribution tar file contains the BIRRP sources, a Makefile, and a pdf version of this documentation. The Makefile may require slight modification for proper compiler optimization. To compile, type “make” or “make all”. To clear all binaries which change when the parameters in the file ‘parameters.h’ are altered, type “make newparameters”. To remove all binaries, type “make clean”.

There are several parameters contained in a file parameters.h which probably will require modification depending on the size of the problem being analyzed. These are:

INTER--logical variable, .true. for interactive operation, .false. for runfile operation. The latter suppresses printing of input prompts on the terminal.

NPCSM--maximum number of separate pieces to each time series

NPTSM--maximum number of data points for all pieces of each time series

NPTSSM--maximum number of values per fft section

NOUTM--maximum number of output time series, normally 3 if two electric and the vertical magnetic field are being processed, but can be 1, 2, or 3

NINPM--maximum number of input time series (i.e., magnetic field), normally 2 but may be 3

NREFM--maximum number of reference time series, must be >=NINPM, and can be much larger than NINPM

NRSITEM--maximum number of reference sites

NSECTM--maximum number of sections, should be set greater than NPTSM*NSCTINC**(NSCTMAX-1)/(NPTSM*OFFS), where OFFS=0.57-0.07*TBW, and NSCTINC, NSCTMAX, and TBW are defined below.

There are additional parameters defined at the top of the main program which should need to be changed only rarely. These may also appear in subroutines as noted in the accompanying comments. These are:

NBLK-- block size for coherence thresholding (set to10)

NFSM--maximum number of frequencies per section to be processed (set to 2)

NFM--maximum number of output frequencies (set to 100)

NFILM--maximum number of filter coefficients (set to 100)

NARM--maximum autoregressive prewhitening filter length (set to 100)

NRECL--maximum block size for binary input (set to 30000)

In addition, functions I1MACH, R1MACH and D1MACH, which provide machine constants, may need to be modified for your computer, although the code contained in version 5 is supposed

to auto-adapt to accommodate either IEEE big- or little-endian arithmetic (among many others).

Time series data are read in by subroutine INPUT which currently supports two formats for the data: 1) formatted data, one value per line, and 2) unformatted data, one or multiple values input per read operation. User specific formats can be implemented by simply changing subroutine INPUT. BIRRP version 2 and later also supports data input from a single ascii file containing all of the data components at one or more sites on sequential lines through subroutine BINPUT. The variables on each line may have any order, and a standard TS format header block can be accommodated.

The program is designed for interactive operation but can easily be run from a shell script, GUI, or command file as well. For the latter, the logical parameter INTER in parameters.h can be set .false. to disable terminal output. In interactive mode, user input is always prompted by a question that is usually self-explanatory. With BIRRP version 2, the program can be run in either basic or advanced mode. Basic mode requires a minimum of input from the user who wishes to compute bounded influence responses with a remote reference. It does not allow 2 stage processing or a variety of other advanced options. Advanced mode is for users who are familiar with bounded influence processing and BIRRP, and allows complete control of the code.

The complete list of required parameters in the order requested are given below. These are color coded in black for basic and blue for advanced mode. All of the basic mode input is required in advanced mode. In addition, some input variables are required or not depending on the responses to earlier ones, and these are noted in the text. Input variables are checked as they are input for compatibility with the dimensions, and warning or error messages are output either on the terminal (INTER=.true.) or in the diagnostic file (INTER=.false.). Fatal errors will result in program termination.

ILEV--selects basic (0) or advanced (1) input mode.

NOUT--the number of output time series (e.g., the electric field for the usual MT response tensor) to be processed. This may either be 1, 2, or 3 in normal use, depending on the size of NOUTM. If NOUT is 1 or 2, then the output time series are both assumed to be horizontal components (such as the electric field). If NOUT is set to 3, the third output time series is assumed to be the vertical magnetic component. This ordering is important when applying rotations later in the

code.

NINP-- the number of input time series (e.g., the horizontal magnetic field for the usual MT response tensor) to be processed. This may be set to 1, 2 or 3, depending on the size of NINPM.

NREF--the number of reference time series. This must be at least as large as NINP, and may be much larger, depending on the size of NREFM, if NRR is set to 1. If NREF>3, then NRR is forced to be 1. If NREF=3, then NRR is forced to be 1 if NINP<3. NREF is required to equal NINP in basic mode, and is not input.

NR3,NR2--the number of reference sites with 3 and 2 components, respectively. This is input only if NREF is greater than 3. Note that each reference site must have at least two data components. Remote reference sites with 3 components must be read in before those with 2 sites.

NRR--selects bounded influence remote reference (0) or two stage bounded influence (1) processing. Note that NREF may not exceed NINP if NRR=0. NRR is required to be 0 in basic mode, and is not input.

TBW--selects the time-bandwidth for the Slepian sequence data taper applied to each section. The time-bandwidth is the product of the data series length and the fundamental bandwidth of the data taper, and must lie in the range 1 to 4. Non-integer values are allowed. Figure 1 shows a set of Slepian sequence data windows at TBWs of 1, 2, 3, and 4. These display increasing time concentration as TBW rises. The bias reduction also increases as the time-bandwidth rises at the cost of increased correlation of adjacent frequencies. Figure 2 shows the corresponding spectral windows (i.e., the square of the Fourier transform) for Figure 1 which display both of these properties. The choice of window determines the overlap between adjacent sections of data in BIRRP (71% for TBW=4, ranging down to 50% for TBW=1). A Slepian sequence with a time-bandwidth of 1 is superior to the cosine bell (Hanning window) in bias protection with similar resolution characteristics. A good choice for TBW is 2 to 3. A time-bandwidth of 4 may be useful for exploratory analysis because no prewhitening is usually needed as it offers excellent bias protection. The choice of TBW will influence the use of prewhitening (essential if

TBW<2; see NAR) and the ensuing correlation properties will influence the frequency increment required to ensure independence of the estimates (see NFINC).

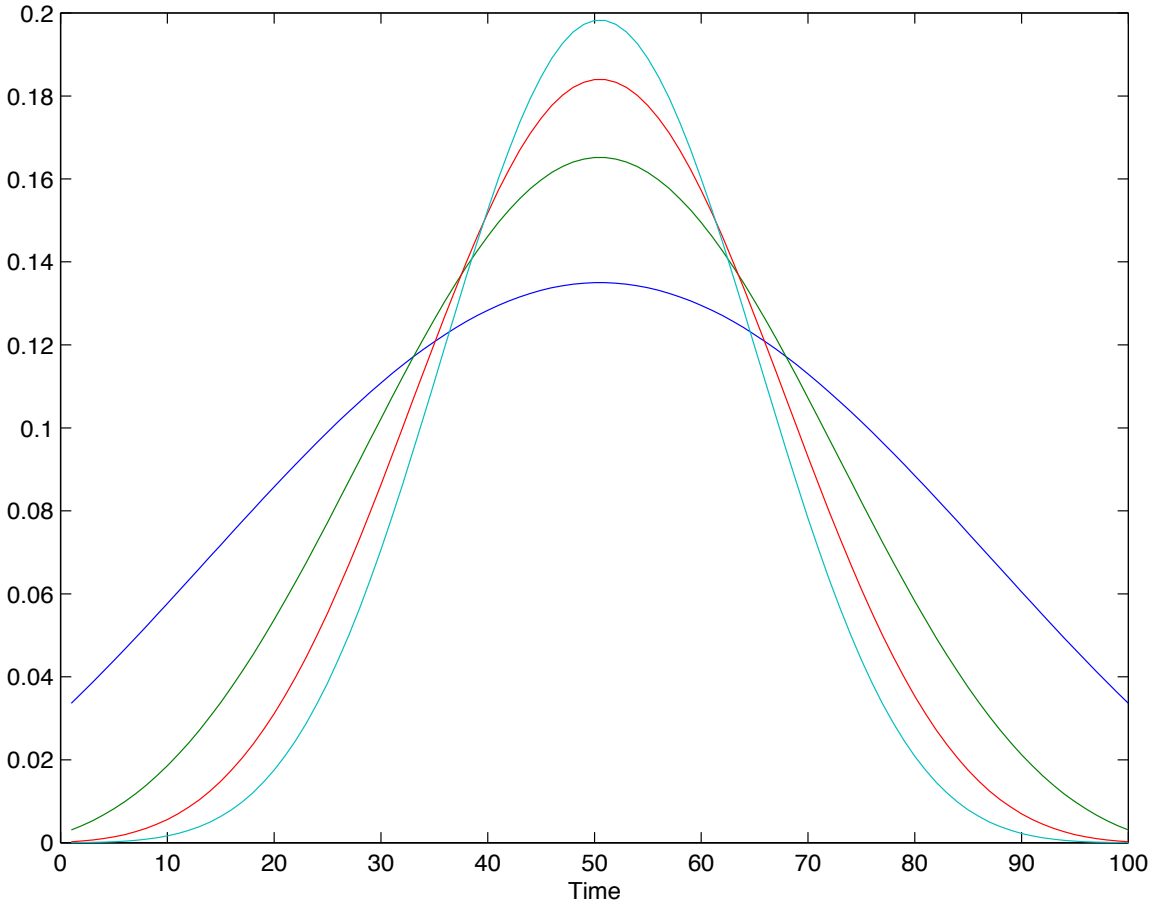

Figure 1. 100 point Slepian sequence data windows with a time-bandwidth of 1 (blue), 2 (green), 3 (red), and 4 (aqua). Note the increase in time concentration as TBW rises, requiring more overlap between data sections to maintain statistical efficiency

DELTAT--the sample interval in dimensional units which should normally be seconds so that the units of apparent resistivity will be correct. If DELTAT<0, then its absolute value is the sample rate, which should normally be in Hz. If DELTAT=0, the sample interval is assumed to be 1.

NFFT,NSCTINC,NSCTMAX--selects the maximum section length (NFFT), factor by which this value is divided when the section length is reduced (NSCTINC), and number of times this operation is performed (NSCTMAX). NSCTINC is set to 2 in basic mode. Thus, the initial section ffts are computed using NFFT, the second pass uses a length of NFFT/NSCTINC, the third pass uses a section length of NFFT/NSCTINC**2, etc., and a total of NSCTMAX section

lengths are used. NFFT is forced to be even if input as odd.

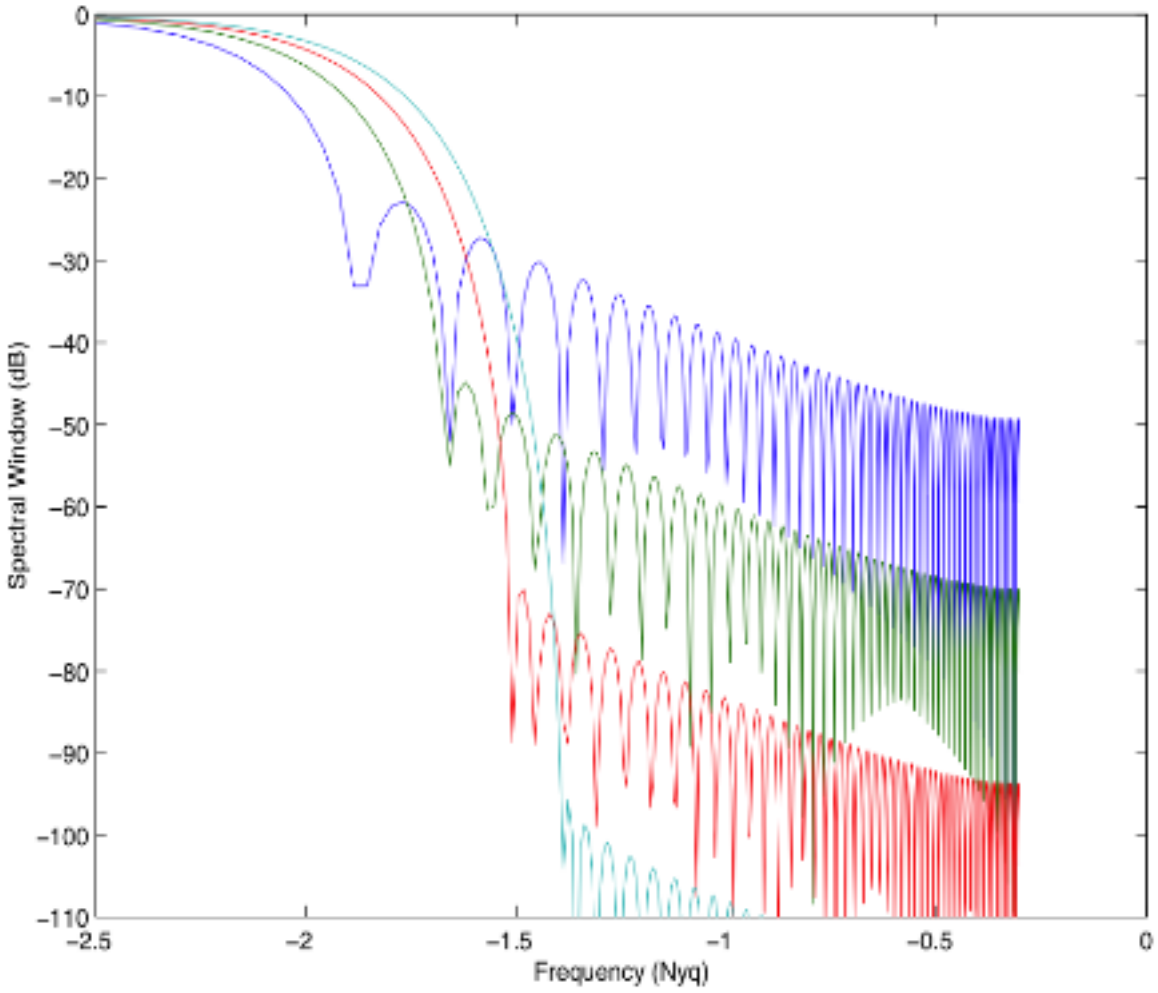

Figure 2. Spectral windows for the Slepian sequence windows of Figure 1 plotted on a log-log scale. For a 100 point sequence, the fundamental (Rayleigh resolution) frequency increment is 1/100 or -2 on the x-axis. The TBW 1 window very nearly has this resolution, and hence adjacent frequencies will essentially be uncorrelated. As TBW rises, the width of the spectral window increases so that only every third frequency will be uncorrelated at a TBW of 4 and approximately every second frequency for TBW 2 or 3. At the same time, bias protection rises with TBW.

NF1,NFINC,NFSECT--selects the index of the first frequency in each section fft (NF1), the increment for subsequent frequencies (NFINC), and the number of frequencies (NFSECT) to be processed. The frequency in Nyquists corresponding to NF1 is NF1/NFFT, where NFFT is the

current section length. Due to possible bias by data window weighting, it is recommended that NF1 be 3 or larger. If a data window with TBW>1 is used, then NFINC should be set to a value large enough that data window-induced correlation will be small. Setting NF1=3, NFINC=1, and NFSECT=2 would select the third and fourth frequencies in each section for bounded processing, which works well if NSCTINC=2 and TBW=1. For TBW of 2 or 3, NFINC should be set to 2 or more, while for a TBW of 4, NFINC should be at least 3. In basic mode, NF1=NINT(TBW+2), NFINC=NINT(TBW), and NFSECT=2, where NINT denotes the nearest integer.

At this point, in interactive mode (INTER=.true.), the code will compute and print out the periods that have been selected by the choice of NFFT, NSCTINC, NSCTMAX, NF1, NFINC, and NFSECT, with all of the values for a given section length appearing on a single line. The user is then prompted to approve this choice with a y or n response. A n response will result in a prompt to re-input the six variables. This feature does not function in batch mode (INTER=.false.). Note that the periods do not need to be monotonically decreasing, as the results are sorted into descending order before output.

MFFT--the factor by which NFFT is divided in computing the robust AR filter. Default value in basic mode is 2, although smaller values may work better for severely contaminated data sets.

UIN, AINLIN, AINUIN--selects the fudge factor for large residual rejection for the robust and small/large leverage point controlled leverage processing components, respectively. If any of these parameters is set to -999, then the respective feature is disabled. AINLIN is automatically set to -999 in basic input mode. If UIN=-999, only conventional processing will be done, and no bounded influence output files will exist. If UIN is set to 0, then the factor is determined automatically from the N-th quantile of the Rayleigh distribution as described in Chave et al. [1987] (this is strongly recommended). If the parameter is larger than 1, then the (N-UIN)-th quantile of the distribution is used, and the downweighting of outliers will become more severe. If UIN lies between 0 and 1, then UIN is interpreted as a probability and the corresponding point on the inverse Rayleigh distribution is used for outlier rejection. If UIN is less than 0, then the factor is given by its absolute value. This must be done with care, else considerable loss of good data and instability of the result may ensue. UIN is always interpreted as

positive in basic input mode, except for a value of -999. If AINLIN/AINUIN lie between 0 and 1, then the value is interpreted as the probability level below/above which leverage points will be rejected. Typical values for AINUIN will lie between 0.95 and 0.9999, with the higher value a good starting point. If used, AINLIN will usually lie in the range 0.0001 to 0.05. If either parameter is larger than one or negative, then the corresponding absolute value is the rejection point. For example, if AINUIN=-2, then leverage values larger than 2 times the expected value of the hat matrix diagonal will be rejected. In basic mode, AINUIN will automatically be set to 0.999 if it is input outside the range (0,1).

C2THRESHB--Selects the squared coherence threshold value for the initial magnetic field processing step of two stage bounded influence processing. Input only in advanced mode if NRR=1. If C2THRESHB is nonzero, the conventional squared coherence between the local magnetic field and that predicted by a linear relationship with the reference magnetic fields is computed by blocks (see NBLK), and the entire block of data is discarded if the coherence is smaller than C2THRESHB. Set C2THRESHB to zero if coherence thresholding is not desired.

C2THRESHE--Selects the squared coherence threshold value for the electric to magnetic field step in either remote reference or two stage processing. If C2THRESHE is nonzero, the conventional squared coherence between the measured electric field and that predicted by a linear remote reference relationship with the local magnetic field is computed by blocks (see NBLK), and the entire block of data is discarded if the coherence is smaller than C2THRESHE. Set C2THRESHE to zero if coherence thresholding is not desired.

NZ--If NOUT=3, this selects the thresholding mode for the third input variable (vertical magnetic field). If NZ=0, the vertical magnetic field is thresholded separately from the electric field, and hence the sections used will be different, precluding use of GBDECOMP simultaneously on the MT and vertical magnetic field transfer function. If NZ=1, the vertical magnetic field is thresholded using the electric field coherence. If NZ=2, the electric and magnetic fields are thresholded together. NZ is not input if NOUT<3.

C2THRESHE1--if NOUT=3 and NZ=0, this inputs the squared coherence threshold to be used for the vertical magnetic field.

PERLO, PERHI--Longest and shortest period in dimensional units according to DELTAT over which coherence thresholding is to be applied. This is not asked for if C2THRESHB, C2THRESHE, and C2THRESHE1 are all zero. If PERLO<0 or PERHI<0, its absolute value is taken to be dimensional frequency rather than period. All periods are thresholded in basic mode.

OFIL--a character string that serves as the root for all output filenames. The response function filenames are obtained by appending .IAJ.rf to OFIL, where I identifies the output time series with a number 1, 2, or 3 according to the order with which the data were input, A is an n for nonrobust or an r for bounded influence, and J is a number 1 or 2 which identifies the input time series, again in the order with which they were read in. For two-stage processing, the stage 1 magnetic transfer function filenames are obtained by appending .IAJ.tf to OFIL, where I identifies the input time series with a number 1, 2, or (rarely) 3 according to the order with which the data were input, A is an n for nonrobust or an r for bounded influence, and J is a number 1, 2,... which identifies the reference time series, again in the order with which they were read in.Files ending in rp contain second stage apparent resistivity and phase using the same naming convention. A single file OFIL.j contains all of the second stage responses in j-format (see http://www.mtnet.info), including a header block containing all of the processing parameters. The prediction coherence filenames are obtained by appending .IA.Jc2 to OFIL, where A is n or r as the estimate is nonrobust or bounded influence, I corresponds to the output (stage 2) or input (stage 1) time series in the order input, and J is either 1 or 2 for stage 1 or stage 2. Filenames for weight files are obtained by appending I.wt.J to OFIL, where I corresponds to the number of the output time series and J is the frequency index. Filenames for qq plots and diagnostics are obtained by appending .I.A.J to OFIL, where I corresponds to the number of the output time series as above, A is oq1, oq2, fq1 or fq2 for original and final qq at stage 1 or 2, or or1, or2, fr1, or fr2 for original and final diagnostics at stages 1 or 2, respectively, and J is the frequency index. Finally, filenames for the bounded influence stage 2 delete one transfer function estimates are obtained by appending .Ir.d1rf to OFIL, where I indicates the index of the output time series. For example, OFIL.1r1.rf would be the Zxx component of the MT response and OFIL.1.wt.2 would be the weight file for the first (usually Ex) output times series at the second frequency.

NLEV--defines the output level as NLEV=-3 to 3. Negative values enable output of the delete-one file, while positive values disable this feature. If NLEV=0, only the response functions are output. This is the recommended setting for most users. If ABS(NLEV)=1, the response functions and qq-plot files are output. If ABS(NLEV)=2, the response functions, qq-plot files, and weight files are output. If ABS(NLEV)=3, the response functions, qq-plot files, weight files, and diagnostic files are output.

NPREJ--indicates the number of frequencies in the raw section Fourier transforms which are to be rejected. This allows a simple form of band reject filtering to be performed, and is very useful in dealing with the solar daily variation and its harmonics.

PREJ--if NPREJ is nonzero, PREJ contains the dimensional periods in nominal seconds in the raw Fourier transforms to be rejected, one value per line. If an entry in PREJ agrees with the actual period to four significant figures, it will be eliminated from processing. If PREJ<0, then its absolute value is taken to be the frequency in nominal Hz. May be input in any order.

NPCS--the number of independent data pieces to be processed for all of the time series. This allows data with gaps to be handled either by storing contiguous sections in separate files or by skipping data sections on input. For example, if there is a missing day in the middle of all of the time series, then NPCS would be set to 2 and the first and second halves of the data would be input separately.

NAR--the order (with leading 1 suppressed) of the robust autoregressive prewhitening filter to be applied to the data. This is recommended if the time-bandwidth is less than 2, and is not harmful at larger values of TBW. If NAR=0, then no prewhitening is applied. If NAR<0, then the data are assumed to already be prewhitened with an AR filter which is input either from a separate file (IMODE=0 or 1) or contained at the head of the data file (IMODE=2). The value of NAR is usually noncritical, and typically will be 3-15, although larger values may give better prewhitening with some data sets.

IMODE--selects the file mode. If IMODE=0, then the data are in separate ascii files, one value per line. If IMODE=1, then the data are in separate binary files with one or more values input per

read operation. If IMODE=2, then the data for one or more sites are assumed to be in a single headerless (except possibly for ar filters) ascii file with each line containing all of the data components at a given time. If IMODE=3, then an ascii file in TS format is assumed.

JMODE--selects the input mode. If JMODE=0, then the number of data to be read for each piece of data and the number of values to skip for each time series will be input in the order specified below. The user is responsible for time series alignment in this mode. If JMODE=1, then the data file start time and data window start and end times for each time series are input in the format YYYY-MM-DD HH:MM:SS. In this mode, the program will compute the appropriate number of values to skip and read for each time series to make the windows coincident for each piece

At this point the time series are to be input in the order output, input, and remote references. If IMODE<2, the hierarchy is piece followed by component (note that this is reversed from BIRRP Version 4 and earlier), so that all of the components of a given piece are read in sequentially. The order of parameter input is reversed for IMODE=2: all of the pieces for a given component are input sequentially for the local and then the remote sites.

For IMODE=0 or 1, there are either 1, 2, or 3 output time series depending on the value of NOUT, and there are NINP input time series. In basic mode or if NRR=0, then NREF will automatically equal NINP, while if NRR=1, NREF may exceed NINP. Note that single station processing can be done in basic mode or with NRR=0 by simply reading the input time series twice at the very slight cost of redundant computation of the ffts. The parameters to be read in NPCS times for NSER=NOUT+2*NINP components in basic mode or NSER=NOUT+NINP+NREF components in advanced mode are:

NREAD--the number of data values to be read for a given piece of data. This is input only once per piece and only if JMODE=0.

The following parameters are then input for each time series in a given data piece:

NFIL--if NFIL>0, this gives the number of real filter parameters (e.g., poles, zeroes, or cutoff fre-

quencies) to be input. If NFIL<0, then a filename character string is input. These are mutually exclusive. The filter response is assumed to be multiplicative (the Fourier transform of the data is multiplied by the filter response at each frequency), but if NFIL=-2, the filter response is inverted before use. If NFIL=0, then no filter correction is applied. NFIL is input only for the first piece of data, and is assumed to apply to subsequent pieces.

FPAR--a real vector of filter parameters NFIL long to be passed to function FILTER. The order and meaning of the parameters is determined by the user. Input only if NFIL>0 and only for the first piece of data.

CPAR--a single character string containing a filename for filter parameter input. Input only if NFIL<0 and only for the first piece of data. The filter parameters in the file are assumed to be a leading multiplicative scalar followed by a series of lines containing dimensional frequency, real part, and imaginary part on each line, and are fit with a taut cubic spline for interpolation to the actual processing frequencies in the code.

NBLOCK--the size of the data blocks for unformatted data. If a single value is input per read operation, then NBLOCK=0. This parameter is not input if IMODE=0. Input only for the first piece of data.

ARFILNAM--character variable containing the complete filename (including path, if appropriate) for the AR prewhitening filter, including leading 1. The length of the AR filter may be any value up to NARM, and may be different for different time series. Input only if NAR<0.

FILNAM--character variable containing the complete filename (including path, if appropriate) for the data series.

NSKIP--the number of leading values to skip in FILNAM. NSKIP may be set to 0 if no skipping is desired. Input only if JMODE=0 and hence mutually exclusive with DSTIM, WSTIM, and WETIM.

DSTIM--19 element character variable in the format YYYY-MM-DD HH:MM:SS containing the data series start time. Input only if JMODE=1 and hence mutually exclusive with NSKIP and

NREAD.

WSTIM--19 element character variable in the format YYYY-MM-DD HH:MM:SS containing the processing window start time. Input only if JMODE=1 and hence mutually exclusive with NSKIP and NREAD.

WETIM--19 element character variable in the format YYYY-MM-DD HH:MM:SS containing the processing window end time. Input only if JMODE=1 and hence mutually exclusive with NSKIP and NREAD.

For IMODE=2, the filter coefficients for each time series are input sequentially, then all of the pieces of the local data set are read in sequentially, and then all of the pieces for each remote data set are read in sequentially. If NREF>3, the hierarchy of three component remote sites followed by two component remote sites is followed. If NAR<0, then a header with the first character '#' is assumed to contain the -NAR+1 term AR filter, including the leading 1, in the same order as for the data. The AR filter coefficients are located immediately after a line headed by '#AR_FILTER' and are read with (1x,5g15.7) format.

The filter parameters to be read in NSER=NOUT+2*NINP times in basic mode or NSER=NOUT+NINP+NREF times in advanced mode are:

NFIL--if NFIL>0, this gives the number of real filter parameters (e.g., poles, zeroes, or cutoff frequencies) to be input. If NFIL<0, then a filename character string is to be input. These are mutually exclusive. The filter response is assumed to be multiplicative (the Fourier transform of the data is multiplied by the filter response at each frequency), but if NFIL=-2, the filter response is computed and inverted before application. If NFIL=0, then no filter correction is applied.

FPAR--a real vector of filter parameters NFIL long to be passed to function FILTER. The order and meaning of the parameters is determined by the user. Input only if NFIL>0.

CPAR--a single character string containing a filename for filter parameter input. Input only if

NFIL<0. The filter parameters in the file are assumed to be a leading multiplicative scalar followed by a series of lines containing dimensional frequency, real part, and imaginary part on each, and are fit with a taut cubic spline for interpolation to the actual processing frequencies in the code.

Subsequently, for each local or remote data set, the input variables are:

NCOMP, (INDEX(I), I=1,NOUT+NINP or NREF--gives the number of components per line (NCOMP) and their location (INDEX) using the standard component naming convention. For example, if there are 5 components in the local data file in the order Bx, By, Bz, Ex, and Ey, then NCOMP=5 and INDEX=4,5,3,1,2. If there are 5 components in a remote data file with the same format, and Bx and By are to be used, then NCOMP=5 and INDEX=1,2.

For each piece of a given data set, the following variables must be supplied:

FILNAM--character variable containing the complete filename (including path, if appropriate) for the data

NSKIP--the number of leading values to skip in FILNAM. NSKIP may be set to 0 if no skipping is desired. Input only if JMODE=0.

NPTS--the number of values to read from FILNAM. Input only if JMODE=0.

DSTIM--19 element character variable in the format YYYY-MM-DD HH:MM:SS containing the data series start time. Input only if JMODE=1 and hence mutually exclusive with NSKIP and NREAD.

WSTIM--19 element character variable in the format YYYY-MM-DD HH:MM:SS containing the processing window start time. Input only if JMODE=1 and hence mutually exclusive with NSKIP and NREAD.

WETIM--19 element character variable in the format YYYY-MM-DD HH:MM:SS containing the processing window end time. Input only if JMODE=1 and hence mutually exclusive with

NSKIP and NREAD.

After all of the data have been read, processing begins. For each fft section length, a series of rotations may be performed. The output, input, and remote reference data may be rotated independently, and the data may have been collected in a nonorthogonal coordinate system. The input and remote reference data are always rotated. If NOUT=1, then no rotation is applied to the output data, while if NOUT=3, then the third input component is assumed to be the vertical one and is also not rotated. Similarly, if NINP is 3, the third series is assumed to be vertical and is not rotated. If multiple remote reference sites are used, then the first NR3 of them are assumed to have 3 components, so that the third component within each is not rotated. The rotation parameters for each set of horizontal time series are:

THETA1,THETA2,PHI--vectors of rotation angles in degrees for respectively the first component with respect to geomagnetic north, the second component with respect to geomagnetic north, and the resulting orthogonal coordinate system after applying the first two rotations. Angles are always positive clockwise. If the original data were collected in an orthogonal coordinate system, then THETA1=0, THETA2=90 and PHI determines the final orientation. The first entry in these vectors is read in only if NOUT>=2 and gives the rotations for the output time series. The second entry is always read in and gives the rotation angles for the input variables. There are 1 (basic mode) or NR3+NR2 (advanced mode) remaining entries for the different reference sites.

The output files contain the following information in G15.7 format with a leading blank. The output files for robust response functions and robust apparent resistivity/phase contain a trailing integer*1 variable. Note that period is always given in dimensional units (DELTAT divided by frequency in Nyquists) and frequency is always given in dimensional units (frequency in Nyquists divided by DELTAT).

OFIL.xn.yc2--Contains coherence information from conventional processing at stage y, where x is an integer corresponding to the index of the input (stage 1) or output (stage 2) variable. From left to right, the entries are period, frequency, the squared prediction coherence, the phase of the prediction coherence (which may be nonzero when a remote reference is used), the theoret-

ical zero coherence value at the 95% confidence level, the lower 95% confidence limit on the coherence, the upper 95% confidence limit on the coherence, the equivalent degrees of freedom, and the fraction of the original data retained after coherence thresholding (stage 1) or coherence thresholding and first stage processing (stage 2). The 1c2 files are output only if two stage processing is being used.

OFIL.xr.yc2--Contains coherence information from bounded influence processing at stage y, where x is an integer corresponding to the index of the input (stage 1) or output (stage 2) variable. From left to right, the entries are period, frequency, the squared prediction coherence, the phase of the prediction coherence (which may be nonzero when a remote reference is used), the theoretical zero coherence value at the 95% confidence level, the lower 95% confidence limit on the coherence, the upper 95% confidence limit on the coherence, the equivalent degrees of freedom, the fraction of the original data retained after coherence thresholding (stage 1) or coherence thresholding and first stage (stage 2) processing, and the fraction of the original data retained after final bounded influence weighting. The 1c2 files are output only if two stage processing is being used.

OFIL.xnx.tf-Contains the nonrobust stage 1 transfer functions, where xx are integers corresponding to the indices of the input and reference variables. From left to right, the entries are period, frequency, the real part of the response, the imaginary part of the response, the standard error on both, the equivalent degrees of freedom, and the fraction of the data used in the calculation after coherence thresholding. Note that the error is identical for the real and imaginary parts. Output only if two stage processing is being used.

OFIL.xrx.tf-Contains the bounded influence stage 1 transfer functions, where xx are integers corresponding to the indices of the input and reference variables. From left to right, the entries are period, frequency, the real part of the response, the imaginary part of the response, the standard error on both, the equivalent degrees of freedom, the fraction of the data used in the calculation after coherence thresholding, the fraction of the data used in the calculation after final bounded influence weighting, and an integer convergence flag (0 for proper convergence, 1 for convergence failure). Note that the error is identical for the real and imaginary parts. Output only if two stage processing is being used.

OFIL.xnx.rf-Contains the nonrobust stage 2 response functions, where xx are integers corresponding to the indices of the output and input variables. From left to right, the entries are period, frequency, the real part of the response, the imaginary part of the response, the standard error on both, the equivalent degrees of freedom, and the fraction of the data used in the calculation after coherence thresholding and first stage processing (if applicable). Note that the error is identical for the real and imaginary parts.

OFIL.xrx.rf-Contains the stage 2 bounded influence response functions, where xx are integers corresponding to the indices of the output and input variables. From left to right, the entries are period, frequency, the real part of the response, the imaginary part of the response, the standard error on both, the equivalent degrees of freedom, the fraction of the data used in the calculation after coherence thresholding and first stage processing (if applicable), the fraction of the data used in the calculation after final bounded influence weighting, and an integer convergence flag (0 for proper convergence, 1 for convergence failure in first stage, 2 for convergence failure in second stage). Note that the error is identical for the real and imaginary parts.

OFIL.xnx.rp-Contains the nonrobust apparent resistivity and phase, where xx are integers corresponding to the indices of the output and input variables. From left to right, the entries are period, frequency, apparent resistivity, the approximate standard error on apparent resistivity from the delta method, the phase, the approximate standard error on phase from the delta method, the equivalent degrees of freedom, and the fraction of the data used in the calculation after coherence thresholding and first stage processing (if applicable).

OFIL.xrx.rp-Contains the bounded influence apparent resistivity and phase, where xx are integers corresponding to the indices of the output and input variables. From left to right, the entries are period, frequency, apparent resistivity, the approximate standard error on apparent resistivity from the delta method, the phase, the approximate standard error on phase from the delta method, the equivalent degrees of freedom, the fraction of the data used in the calculation after coherence thresholding and first stage processing (if applicable), the fraction of the data used in the calculation after final bounded influence weighting, and an integer convergence flag (0 for proper convergence, 1 for convergence failure in first stage, 2 for convergence failure in second stage).

If ABS(NLEV)>1, then there will be a series of additional files as discussed earlier. These are all in G15.7 format with a leading blank.

OFIL.xx.xqx.x--data for plotting quantile-quantile plots of the residuals and hat matrix diagonal for either conventional regression or bounded influence regression for either stage 1 or 2. The file contains the period on an initial line followed by 5 zeroes, then (in order) the quantiles, order statistics, and section indices for the regression residuals and the hat matrix diagonal on separate lines (total of six entries per line). All of the q-q data are output at the extremes (upper and lower 10%), while fewer are output in the middle of the distribution. Output if ABS(NLEV)>=1.

OFIL.xx.wt.x--contains the index of the middle of the section in dimensional time, the block coherence estimate(s), the final robust weight, and the final hat matrix weight for each section. The number of block coherence estimates depends on whether two stage processing is used; only one is output in basic mode or if NRR=0. Output if ABS(NLEV)>=2.

OFIL.x.yrx.x--data for plotting a variety of diagnostics for the conventional and bounded influence regression problems for either stage 1 or 2. For the nonrobust version, the file contains period and the maximum residual, output, and input power on an initial line, then the normalized hat matrix diagonal, normalized residual power, and normalized output and input powers for each section. Normalization is relative to the maximum value for the residual, output, and input powers and relative to the expected value for the hat matrix diagonal. For the bounded influence version, the file contains the same entries except that the appropriate weights have been applied to the powers.

A complex external function FILTER must also be supplied by the user so corrections for the system response can be applied if NFIL>0. The call is of the form

FIL=FILTER(FREQ,FPAR,NFIL)

where FIL and FILTER are complex variables, FREQ is the dimensional frequency (i.e., Nyquist frequency divided by DELTAT), and FPAR is a real vector NFIL+1 long containing the sample interval DELTAT in the last position after NFIL real filter parameters. The filter response is assumed to be multiplicative (the Fourier transform of the data is multiplied by FIL at each fre-

quency).