EDGES MEMO #387

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
HAYSTACK OBSERVATORY
WESTFORD, MASSACHUSETTS 01886

February 8, 2022

Telephone: 617-715-5533

To: EDGES Group

From: Alan E.E. Rogers

Subject: Haystack EDGES processing pipeline

The Haystack EDGES processing pipeline was started in 2012 with a brief description in memo 96. An update along with a description of the parameters was made in 2015 in memo 113. An especially critical part is the RFI excision described in memo 244.

There are 3 steps of processing carried out by c-coded programs acqplot, edges2 and longav as described in memo 211. These key processing stages are:

1] Reading the raw "acq" data files, conversion to 3 positioned switched data and RFI filtering

2] Calibration and beam correction of data from acqplot

3] The final stage of RFI filtering, averaging over data blocks of time, GHA or days from a single output file or a concatenated files from edges2 followed by a multi dimension grid search for the 21-cm feature or estimates of signature parameters and confidence using delta Chi-squared boundaries as in memo 272.

The pipeline is run using a c-shell script in which the processing parameters are passed via arguments to the complied c-programs. An example of a script using beam file newniv.txt and calibration file specal.txt is shown below:

```
#!/bin/csh -f
# final case2 lowband1 2016_250 - 2017_095
set c = S11_blade_low_band_2015_342_03_14.txt.csv
cp newniv.txt azelq.txt # used for H2
cp specal_final_case2.txt specal.txt  # calibration data file

set day = 100
while ($day <= 300)

acqplot $day.acq  -rfi 2.5 -tcal 1000 -pfit 37 -fstart 40 -fstop 100 -smooth 8 -pkpwrm 40 -peakpwr 3 -minpwr 0.7
-nrfi 4 -dloadmax 1000 -adcov 0.4 -maxrmsf 200 -maxfm 200

edges2  spe.txt -fstart 50 -fstop 100 -spant spe.txt -s11ant  $c -mfit 3 -wfstart 51 -wfstop 99 -lmode 6 -tant 296 -
eorcen 150 -skymode 384 -antaz 354 -wtmode 100 -nfit4 10 -smooth -8 -low 1 -mdd 4 -delayant  0.0e-12 -adb
0.0 -cmb 2 -ldb 0.0 -delaylna 0e-12 -adb 0.0 -delayant 0.0e-12 -test 0 -eoramp -0.5 -eorwid 0 -eorcen 78 -tau 7

cat spe0.txt >>! spesum.txt
@ day = $day + 1
end
longav spesum.txt  -lim 0.16 -nfit 5  -dmax 0.5  -fstart 53 -fstop 99 -schk 0 -tchk 200 -rfi 2 -g10  -date -seor -1 -
tau 7 -sig 0 -md 1
```

Calibration of the raw VNA S11 data requires an additional c-coded program reads1p4 along with a script to obtain the S11 of the hot load, ambient load, open cable, shorted cable, LNA and antenna along with measurements of calibration SOL. For EDGES-2 the calibration data in taken in the lab along with the S11 data using a benchtop VNA to calibrate the receiver "internal" calibration SOL in the front-end to allow remote antenna S11 measurements from the VNA in the electronics hut. For EDGES-3 the VNA is in the antenna and the entire calibration is fully automated using three 2-position and one 8 position low loss mechanical switches shown in the block diagram in Figure 1a of memo 300.

Test features of the pipeline:

1] Offsets to S11 measurements can be applied during calibration and/or processing to help evaluate the effects of systematic errors in VNA measurements

2] Simulated data can be generated by edges2 to help investigate the effects of systematics and sensitivity to sources of error. This feature is especially useful in the determination of the antenna beam chromaticity of the different FEKO models and different sky maps. Special c-code is available outside of the pipeline to simulate raw data in the acq format which can then be processed by the pipeline. An example is the simulation of the circuit model of the LNA and its noise waves in memo 334.

Details of the arguments of the EDGES processing programs follow:

acqplot program arguments:

input is file or concatenated files from pxspec or fastspec
followed by
-mode default 0 = 3-pos-sw 1 = antenna 2 = load 3 = cal 4,5,6,7 test mode
-fmode default 0 = fourier fit 1 = poly fit
-scaledb default 0 = linear scale > 0 = log scale for plot
-wscalemax default = 1000 = max scale of water fall plot
-wscalemin default = 0 = min scale of water fall plot
-tstart default = 0 = start hour
-tstop default = 24 = stop hour
-pfit default = 0 = no fit > 0 = number of terms
-ns default = -1
-nn default = 256
-d150 default = 1e99 < 1e6 = level of RFI around 150 MHz to be zero weighted
-dloadmax default = 1e99 < 1e6 = level of load for deletion of block
-smooth default = 0 > 0 = number of frequency channels to be averaged in output file spe.txt
-day default = 0 > 0 = day of data to accept
-gal default = 0
-gha default = 0 = Galactic hour angle +/- dgha to accept
-dgha default = 48 +/- range of Galactic hour angle to accept
-nstart default = -1 > -1 3-pos-sw cycle to accept
-tcal default = 400
-fstart default 0 = start frequency MHz
-fstop stop frequency MHz
-peakpwr default = 1e99 level peak power for deletion of blocks
-minpwr default = 0 level minimum power for deletion of blocks
-pkpwrm default = 1e99 level peak power above 80 MHz for deletion of blocks
-sunlim default = 1e99 sun elevation limit for accepted data
-moonlim default = 0 > 0 = accept only data with moon elevation above this limit < 0 = accept only data with moon elevation below this limit
-trec" if present compute receiver temperature

-water default = 0 if not 0 generate waterfall output water.txt and plot in water.pos
-nrfi default = 0 > 0 = number of adjacent channels to be deleted to channel with rfi
-recom default = 0 1 = use only antenna data to avoid added noise in 3-pos-sw data
-adcov default = 0 > 0 = level of adc for deletion of block
-maxscale default = 0 > 0 = maxscale of plot in spe.pos
-minscale default = 0 > 0 = minscale of plot in spe.pos
-mxsweep default = 1e99 = limit of frequency sweep for rfi
-maxrmsf default = 1e99 = rms limit for deletion of blocks
-maxfm default = 1e99 = rms limit for rfi zero weighting of channels in units of sigma in the FM band
-delaystart default = -1e99 = number of seconds delay in the start of data acceptance from the beginning of data file
-wwd default = 0 number of blocks to average before doing 3-pos-sw calculation
-rfi default = 0 rms limit for rfi zero weighting of channels in units of sigma in final average

Outputs: spe.txt - uncalibrated average output spectrum
water.txt - waterfall output
plots: spe.pos – pos files are in postscript format whicg can be “grep”ed
water.pos waterfall plot

edges2.c program arguments:

input file name followed by
-fstart - frequency start MHz needs to be first if used
-fstop - frequency stop MHz needs to be first if used
-spant antenna spectrum file
-spopen
-spshort
-sphot
-spcold
-s11ant antenna s11 file
-s11open
-s11short
-s11lna
-s11hot
-s11cold
-cals11 calibration S11 file with all
-cals11_cold pulls cold load S11 from calibration S11 file
-cals11_hot
-cals11_open
-cals11_short
-cals11_sim1
-cals11_sim2
-cals11_sim3
-cals11_sim4
-cals11_cbox
-cals11_3db
-cals11_cbx
-cals11_noise
-cals11rig – sparms rig format
-hotsparms – hot load sparms
-cals11_sm33 – s11 0f simulator 3
-nocal default 0 1 = nocal
-ydhms time for simulated data

-skymode default = -1 = no beamcorrection 0 = beamcorrection 256 = use beamcorr as data + 1,2,4,8,16,32,64,128 opts
-skymod2 use for new slower beamcorrection using new Haslam map
-lmode default = -1 = to use aloss 0=for hot load 2=hot+rigloss 1= balun + bean using interpolated antloss 5or6=lowband balun +conn
-wtmode default = 0 sets weighting for fit and > 10 for calibration 1 = normal zero wt outside wfstart & wfstop
-nfit1 default = 27 number of terms in fitting to calibration data in specal.txt
-nfit2 default = 27 number of terms in calibration fitting of s11 for hot,cold and cable
-nfit3 default = 27 number of terms in fitting LNA s11
-nfit4 default = 37 number of terms in fitting
-mfit number of parameters in fit to spectrum
-wfit default = 4 number of terms in fitting noise waves
-bfit default = 0 optional fitting of beam data
-cfit number of terms in final calibration fitting not used if specal.txt is used default=7
-smooth optional smoothing of antenna spectrum after fitting
-cons optional fitting constraints
-aloss antenna loss in dB as an alternate simple antenna loss needs lmode = -1 to activate
-atten attenuation between antenna and LNA used in simulation
-adb default = 0 correction to ant s11 in dB
-ldb default = 0 correction to lna s11 in dB
-opn default = 0 correction to lna to open at VNA
-thot temperature of hot calibration load
-tcold temperature of cold (normally room temperature) calibration load
-tcab temperature of cable used for noise wave calibration get set to tcold unless entered
-tant temperature of antenna used in loss calculation needed even when specal.txt is used
-wfstart start of weighted data default=50
-wfstop stop of weighted data default=200
-fbstart optional beamfit start
-fbstop optional beamfit stop
-delaylna default = 0 adapter correction on LNA S11
-delayant default = 0 correction for antenna
-dlyrob not used one-way delay in Roberts balun
-delaycorr default = 0 correction to VNA delay
-dbcorr default = 0 correction to VNA dB
-Lh loss of hot load at 100 MHz uses tamb for loss calc Lh = -1 for hot load model Lh = -2 to use s12,s22
-eorcen eor model center freq 0=no EOR search forces different set of parms
-eorwid eor model FWHM width 0=no EOR search
-eoramp eor model amplitude
-antaz antenna azimuth
-dscale fix plot scale
-specin default = -2.5 assumed foreground spectral index
-sunind default = 0.5 assumed sun spectral index
-binteg beamcorrection time span in secs default 1 used in simulations
-mdd default = 0 = pow(f/150,-2.5+i) mdd=1 for gamma etc.
-sim default = 0 > 0 to simulate data
-test default = 0 test = 1 use beamcorr as data
-cmb default = 0 1 = subtraction of cmb for spectral index and beam correction
-parm1 not used
-parm2 not used
-parm3 not used
-parm4 not used
-ion add ionosphere
-tau signature model opacity

-noise default = 0 add noise
-rr change srand
-site 0 = default = mro 1 = oregon
-map = 0 = default = Haslam 1 = Guzman
Inputs: file from acqplot.c
other files needed in current directory
For beam correction Haslam map 408-all-noh
Calibration file if not running calibration

Output files: spe0.txt
specal.txt from calibration
plots (in ASCII postscript format:
spe0.pos spectrum output
spe5.pos balun loss
spe99.pos calibrated spectrum
spevna0 antenna S11 with calibration
spe1.pos spe2.pos spe3.pos
specvna1.pos spevna2.pos spevna3.pos spevna4.pos spevna5.pos spewav.pos

longav.c program arguments:

Data file or concatenated files from edges2
followed by

-dmax datamax in plot
-lim default = 10 limit above which data with higher rms is not accepted
-fstart frequency start
-fstop frequency stop
-eor default = 0 add eor signature
-aeor
-feor frequency of signature
-seor = -1 to perform a grid search for signature
-wid FWHM of signature
-nfit number of terms
-wtmode - not used
-imode - fit ionosphere parameters
-pmode default 0 plotmodes 1 = no average in spe.pos
-out default = 0 1 = generate output file spe0.txt
-sig default = 0 = physical terms 10 = polynomial terms
-test
-tau default = 0 = Gaussian > 0 flatenning parameter
-adnoise default = 0 adnoise = added random noise amplitude
-rfi rms limit for spectrum accepted into average
-mchk default = 0 1 = accept spectrum if moon is below horizon 2 = accept spectrum if moon is above horizon
-tchk default = -1e99 only accept spectrum if temperature is above tchk
-schk default = 0 1 = reject spectrum if sun is above horizon
-g10 reduce plot scale for average by factor of 10
-gg100 reduce plot scale for average by factor of 100
-diff default = 0 > 0 obtain difference spectrum
-rr srand value to generate new random noise sequence
-alt difference between alternate spectra
-half difference between first and second half
-sim default = 0 > 0 simulate data with scale of sim

-resid default = 0 1 = fit to residual spectrum
-sm default = 0 1 = fit to residual spectrum
-md default = 0 1 = plot signature + residuals in spez.pos
-ti specify keyword for information in spe.pos
-date use date for information in spe.pos
-datt - not used

Inputs: file from edges2 which can be
edges2 outputs for each day concatenated into one large file

Output files: spe0.txt output file
spez.txt signature search output file
Plots: spe.pos output plot
spez.pos signature search plot

reads1p4.c program arguments:

-Tfopen remote OSL VNA measurements
-Tfshort
-Tfload
-Tfant remote antenna VNA measurement
-Tlopen external OSL VNA measurements
-Tlshort
-Tlload
-Tcopen remote VNA measurements of internal OSL
-Tcshort
-Tcload
-res resistance of external standard load
-loadps delay of external standard load
-openps delay of external standard open
-shortps delay of external standard short
-lna sets mode used in measurement of LNA S11

Inputs: raw VNA s11.s1p files
Output files: s11.csv - corrected s11 file in csv format

This program is used in conjunction with the C-shell script
dos1p to obtain the measurements on the simulator hot cold open and shorted cable S11
values when attached to the receiver during calibration and to obtain calibrated
measurements of the LNA S11 when the VNA is connected to the receiver input.