gDAS IP time series data processing

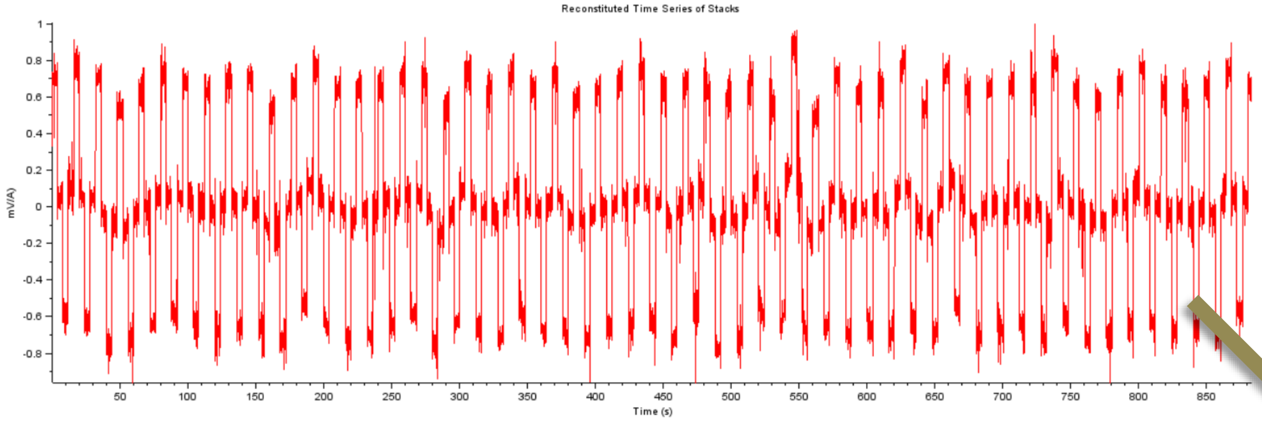
ASEG IP Workshop 2016

IP processing and QC - from amps in the ground to an Inversion input.

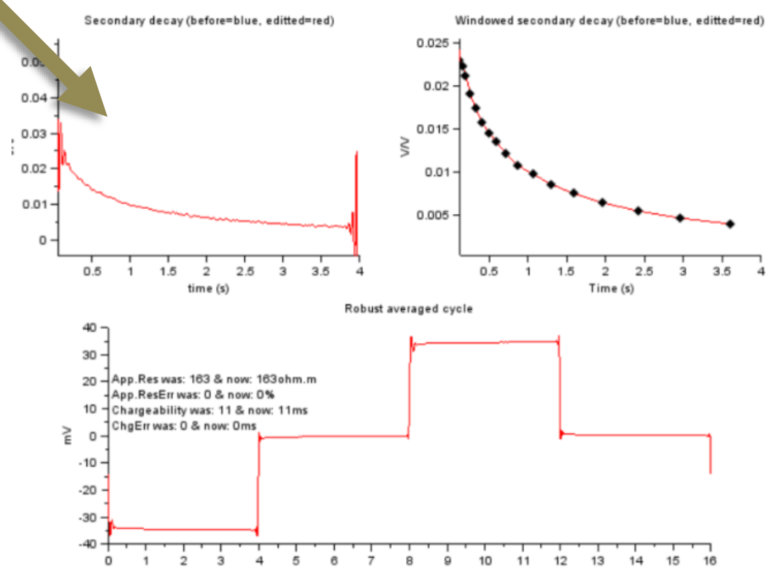
Presented by:
Jeremy Barrett
General Manager

The aim of time series data processing

…. is to enhance the relevant part of a recorded signal, or conversely, diminish the unwanted part.



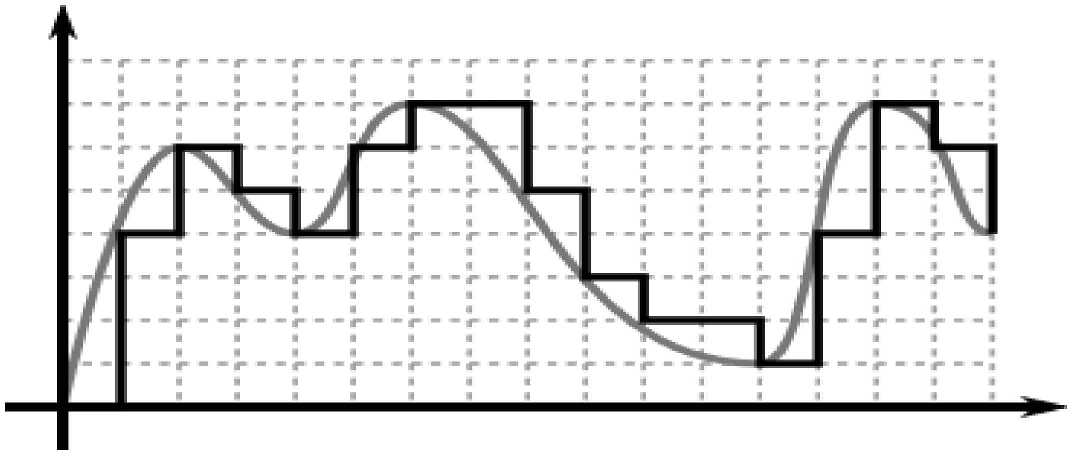
Raw observed E-field time series data



Telluric Cancelled, filtered, editted, stacked and averaged IP data

The first stage of data processing

- The true nature of the signals that we measure is an analogue, constantly variable, signal
- All receiver systems filter and then digitise the true signal to generate a time series of essentially discrete measurements separated in time usually at a constant (or several different) sample rates.
  - As such all receiver systems carry out some level of data processing prior to recording the time series data.



Acquisition level processing

Optimal analogue (pre-digitization) filtering carried out by the receiver system depends on the:

- Characteristics of the signal (and noise)
- Characteristics of the ADC
- The approach to digital data processing

The approach has been for minimal analogue pre-digitization filtering such that the recorded time series is the most faithful representation possible of the signal received at the input terminals.

In extreme conditions (very strong cultural noise for example) additional pre-ADC filtering can however be beneficial.

Why use minimal analogue filtering ?

Digital filters have certain advantages over their analog counterparts:

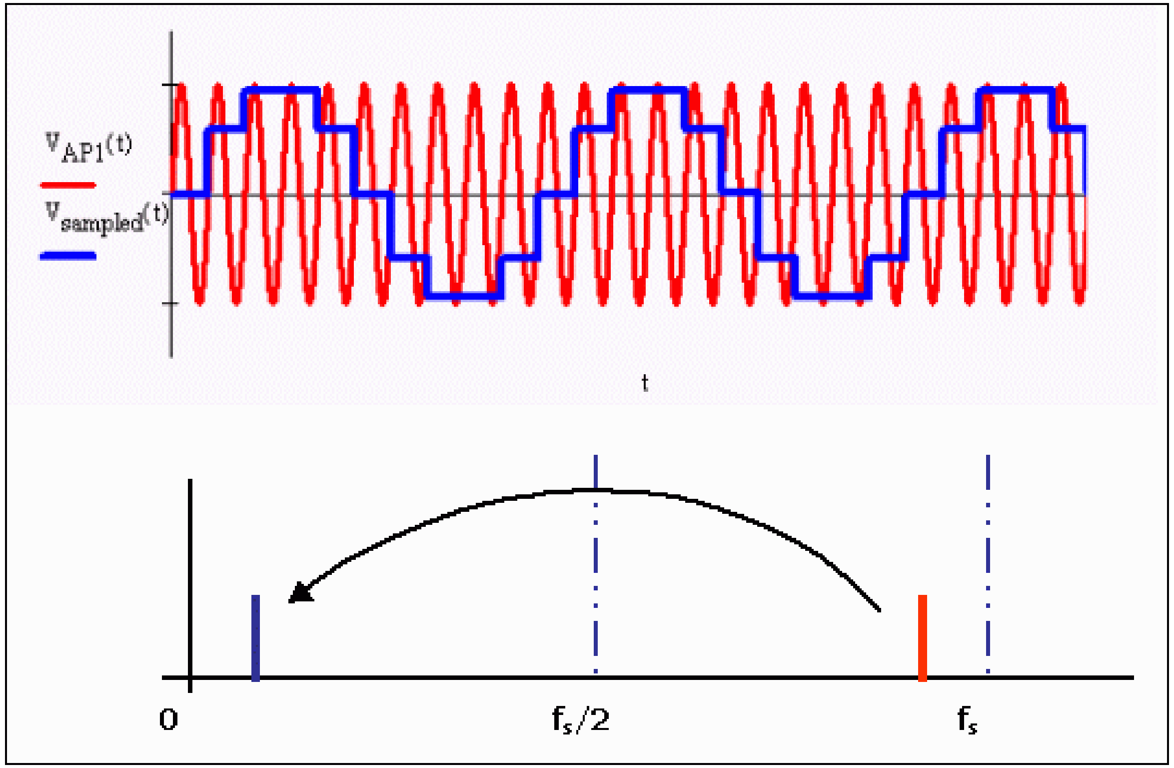
- They are programmable in terms of filter order, cutoff frequencies, and amount of ripple
- They are stable and predictable
    - They don’t drift with temperature or humidity
- They don’t require precision components

- **And, after acquisition data can be re-filtered in different ways**

Anti-alias filters

As an absolute minimum it is necessary to use an antialias filter prior to digitization.

An antialias filter is a low-pass filter set at some frequency less than the Nyquist frequency such as to prevent (limit) higher frequency signals becoming aliased in the digitization process.



Sampling rate

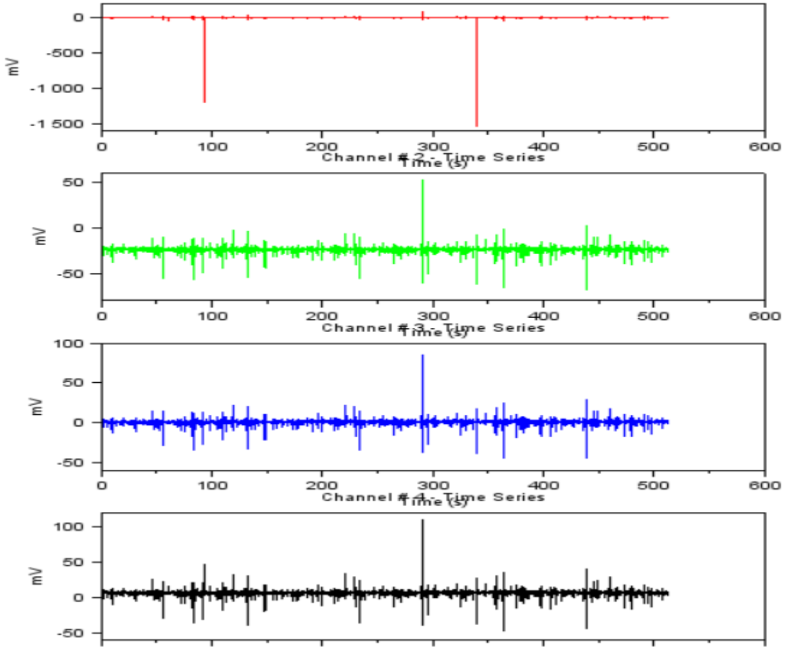
Sample frequency (F_s) controls the highest frequency of the bandwidth that can be analysed, at most given by the **Nyquist frequency ($F_s/2$)** but often in practice limited to half-Nyquist ($F_s/4$).

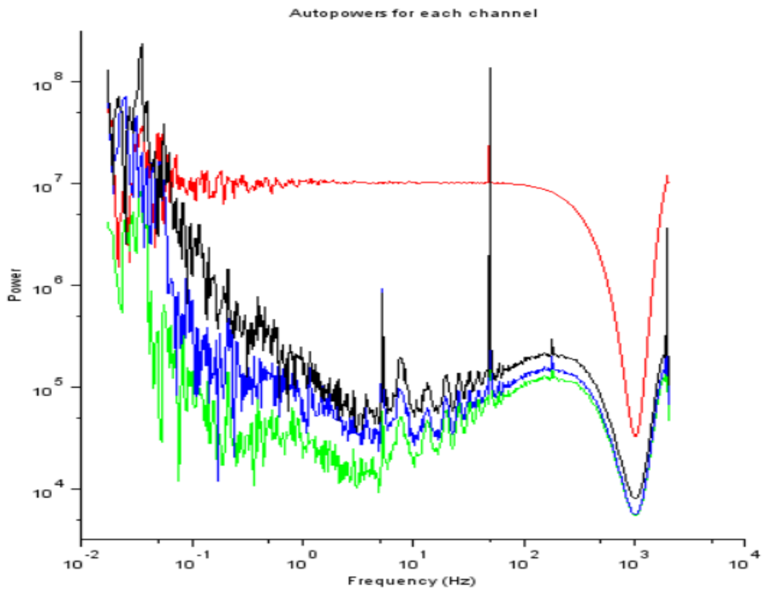
Typical gDAS F_s: 32, 256, 2048, 16384Hz

With F_s = 256Hz usually used for gDAS IP acquisition

Initial stages of time series processing

- Detection and removal of saturated data
- Detection and removal of “spikes”

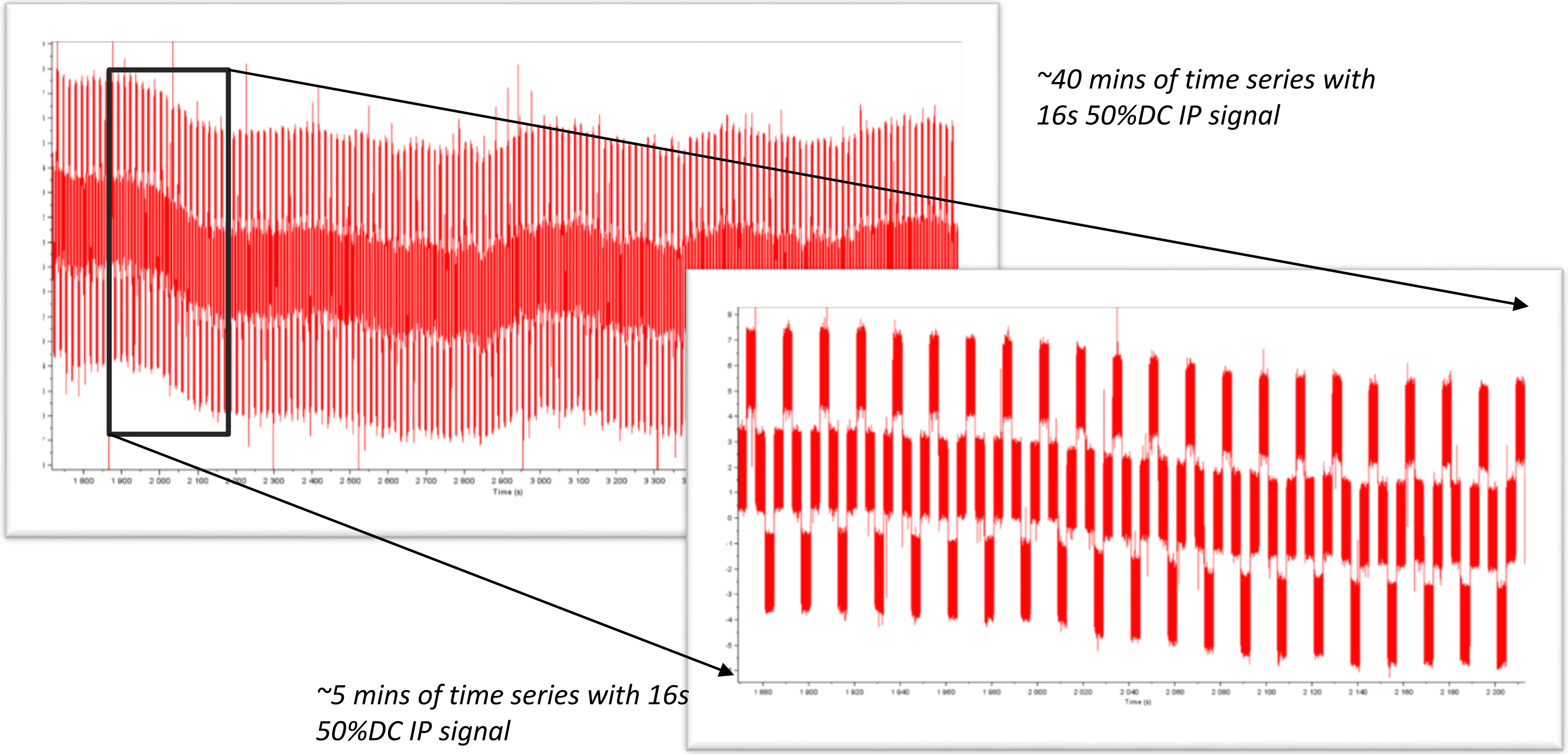




- Replacement of removed data points with inferred data
  - Methods:
    - Assign with zero value
    - Interpolate (linear, low-pass filter, or spline)
    - Interpolate with synthetic data based on signal within rest of the dataset (eg. underlying 50Hz power-line signal)

Time series data from IP acquisitions

The (IP) signal of interest is mounted on long period telluric and perhaps electrode “noise”, with other higher frequency spheric, cultural and other noise sources adding further complexity to the signal.



~40 mins of time series with 16s 50%DC IP signal

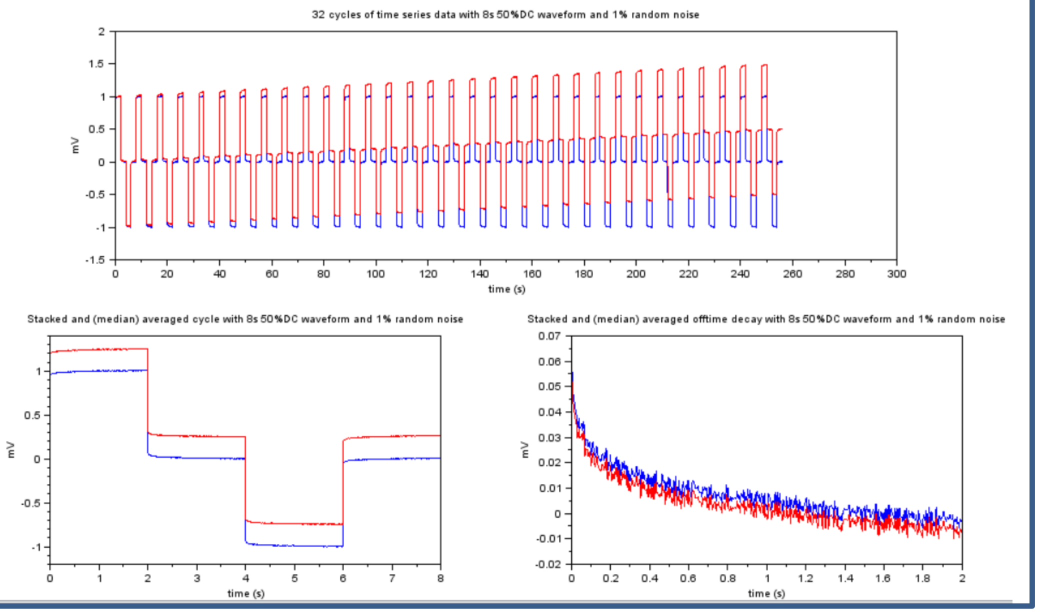
~5 mins of time series with 16s 50%DC IP signal

Obtaining a baseline for the time series

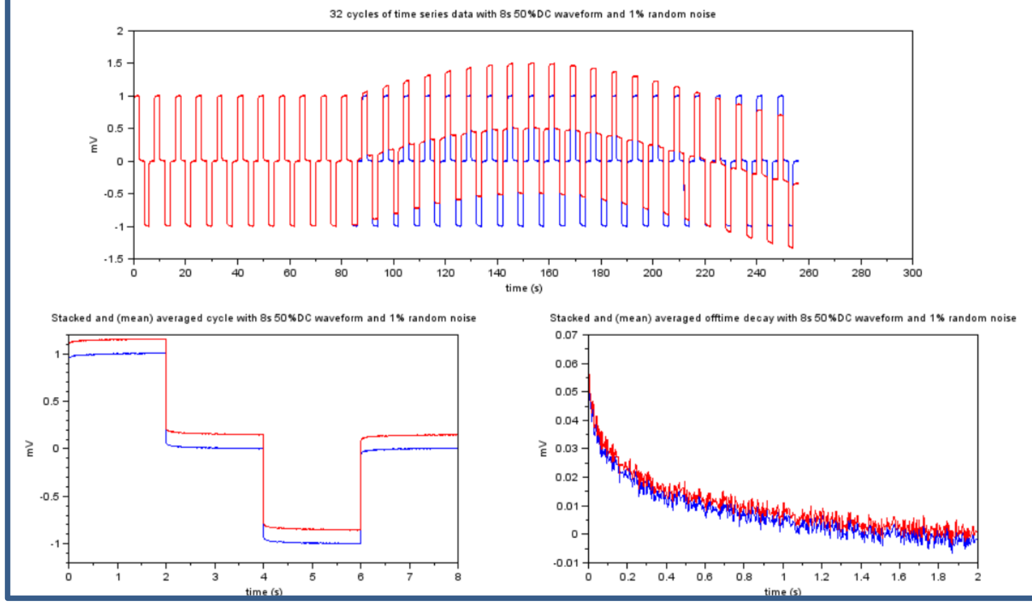
Inferring a baseline for removal from the time series trace is fundamental.

As a rather simplistic example, an uncorrected constant, non-zero, drift if simply stacked (two half-period filter) would bias the calculated chargeability.

M=4.9, M_{biased}=1.2ms



M=4.9, M_{biased}=7.7ms



Methods of estimating the baseline

- Fit a linear trend
    - Halverson stacking incorporates this into the stacking algorithm which is based on values of ¼, -½ , ¼ on three consecutive half-cycles as opposed to values of ½, -½ on two consecutive half-cycles.
- Low pass filtering
- Fitting a curve - Assign the average value for (part of) each complete cycle to the central time of that cycle and fit a curve (monotone spline for example) to these points
- Moving average filter
    - Boxcar of the same width in time as the transmitter cycle

The main problem is that these methods cannot remove noise at or close to the fundamental and higher odd harmonics of the transmitter (source) waveform without distorting the signal of interest.

Method for inferring the baseline Telluric Cancellation

Inference of the **telluric** component of the observed signal is achieved by **measuring representative magnetic field variations at a remote site** and with **knowledge of the transfer function of the Earth at the local measurement site** the corresponding (telluric) electric field may be inferred.

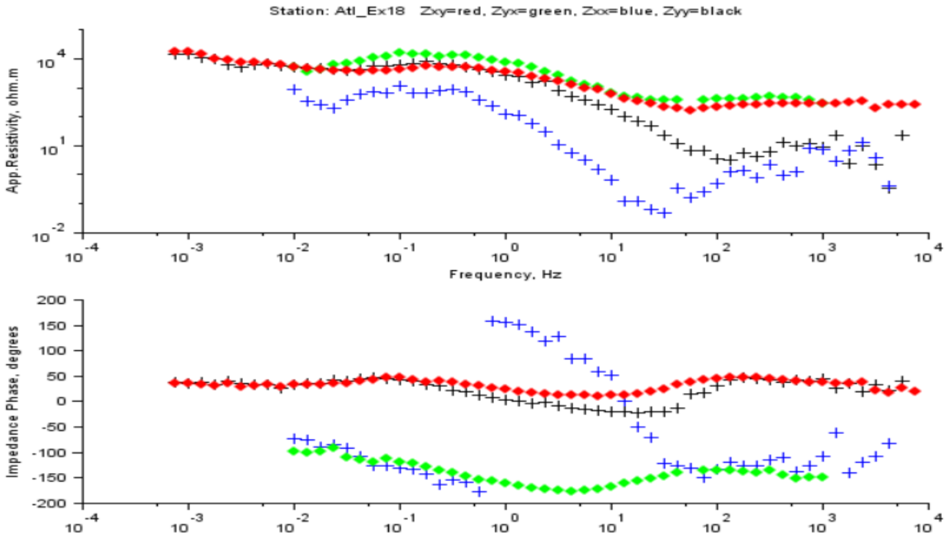
The transfer function in this case refers to a continuous function that relates the Electric [E] to the Magnetic [H] field variations, in MT called the Impedance tensor, [Z].

- [Z] is obtained at discrete frequencies by acquisition of Magneto-Telluric (MT) data at the receiver dipole (E-field) site of interest in the absense of a controlled source, where:

$$[Z] = \frac{[E]}{[H]}$$

- [Z] is a tensor,

$$\begin{bmatrix} Zxx & Zxy \\ Zyx & Zyy \end{bmatrix}$$



Because [Z] changes smoothly as a function of frequency, it is fairly trivial to interpolate to all frequencies required in the bandwidth of interest.

Inferrence and subtraction of telluric noise
Telluric Cancellation

So we look to calculate $[E_i] = [Z][H_{RR}]$ during current injection

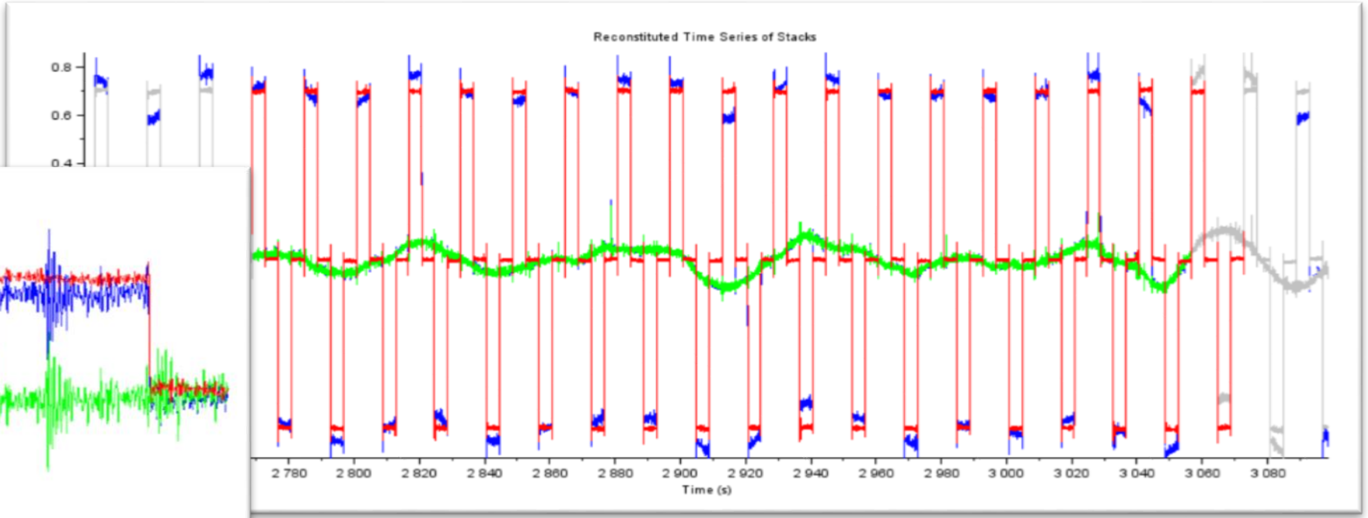
- $[E_i]$ is the inferred telluric electric field at the local site

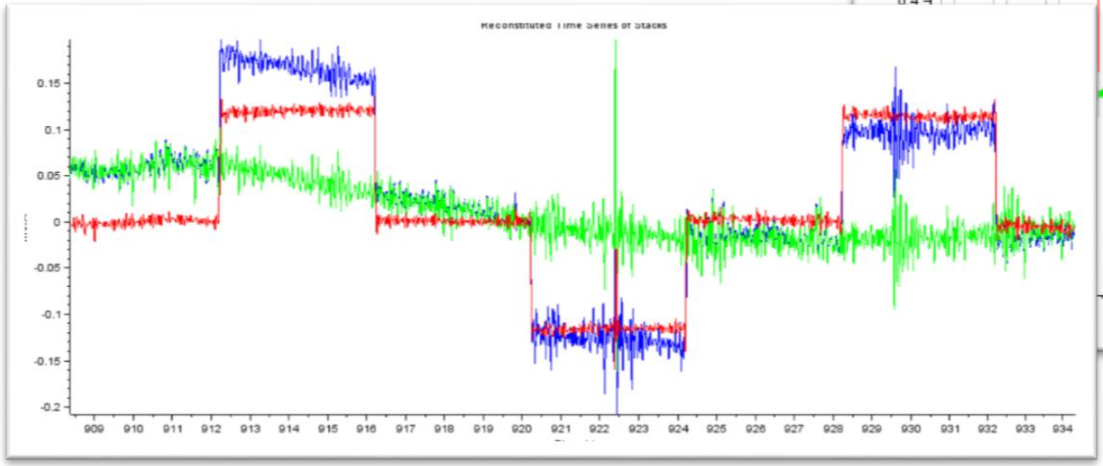
The Telluric Cancelled electric field is then simply,

- $[E_{TC}] = [E_{obs}] - [E_i]$

This approach will go wrong if:

- $[H_{RR}]$ at the remote site is affected by the source signal used for the IP measurement
- $[H_{RR}]$ at the remote site is not the same as at the local E-field site
- $[Z]$ or $[H]$ is erroneous



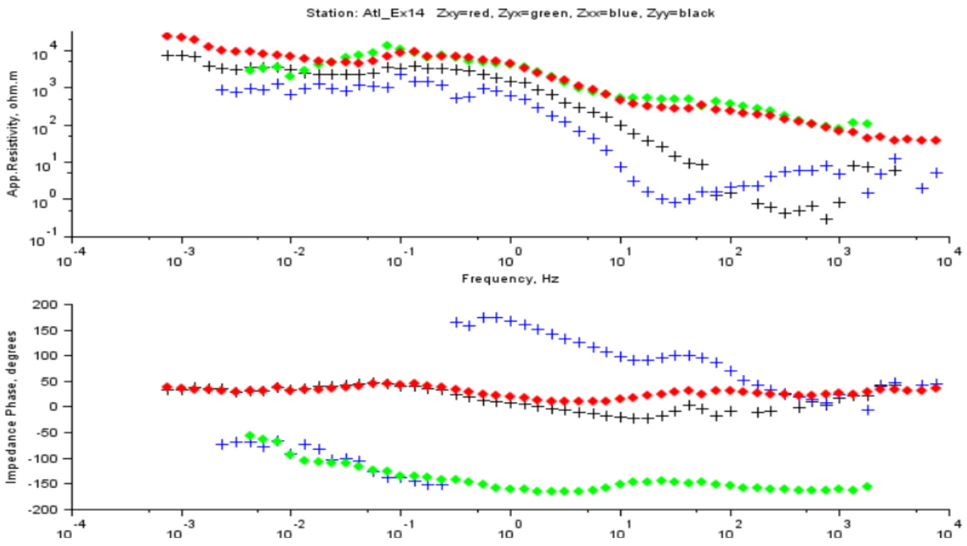


Telluric Cancellation
Selection of [Z]

The use of calculated impedance responses from 1D or even 2D inversion of MT data may be used as the transfer function for telluric cancellation.

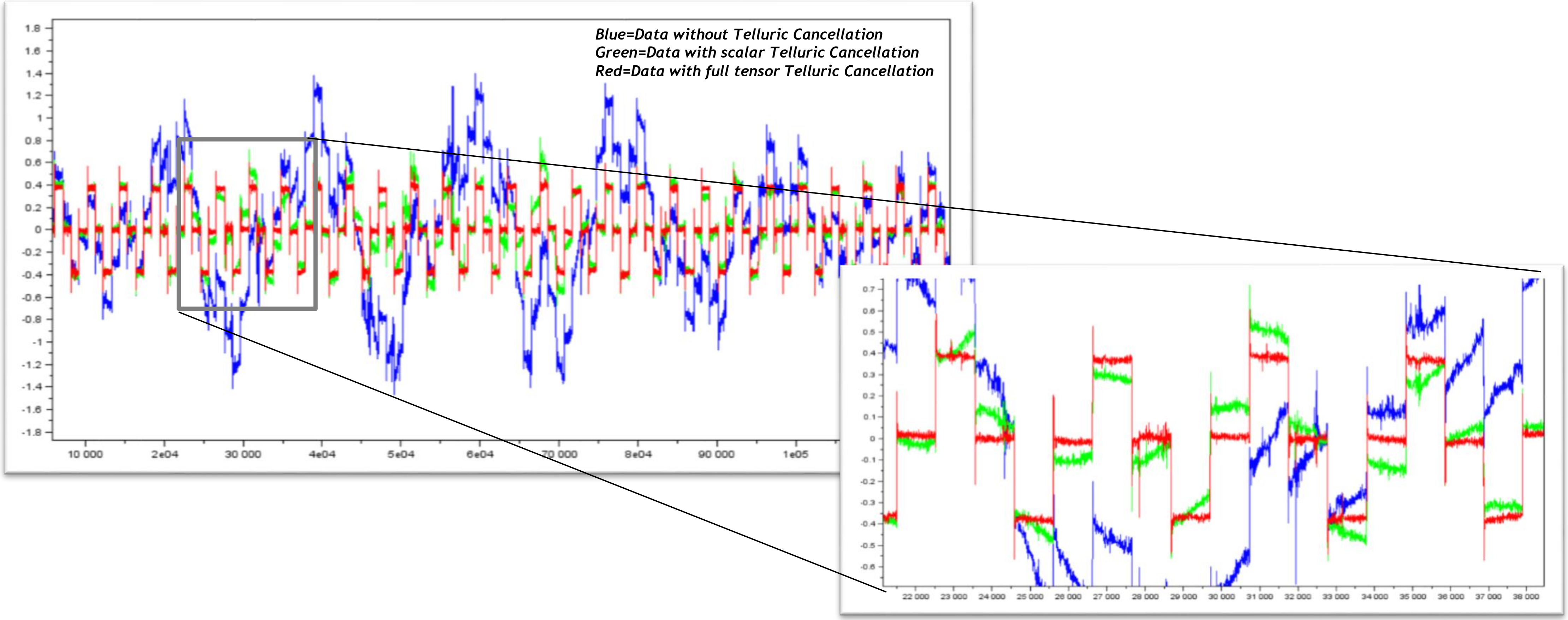
- Benefit: Obliges the transfer function (Z) to be smoothly varying
- Deficiency: The model fit may be poor, but most importantly the principal diagonal elements of [Z], ie. Z_{xx} and Z_{yy}, are by definition zero in these models.
  - This is often an adequate assumption in the ~1D near surface and hence high frequency part of the spectrum but often inappropriate at long period (deep) portions of the spectrum.

Use of (observed) full tensor impedance is preferred.



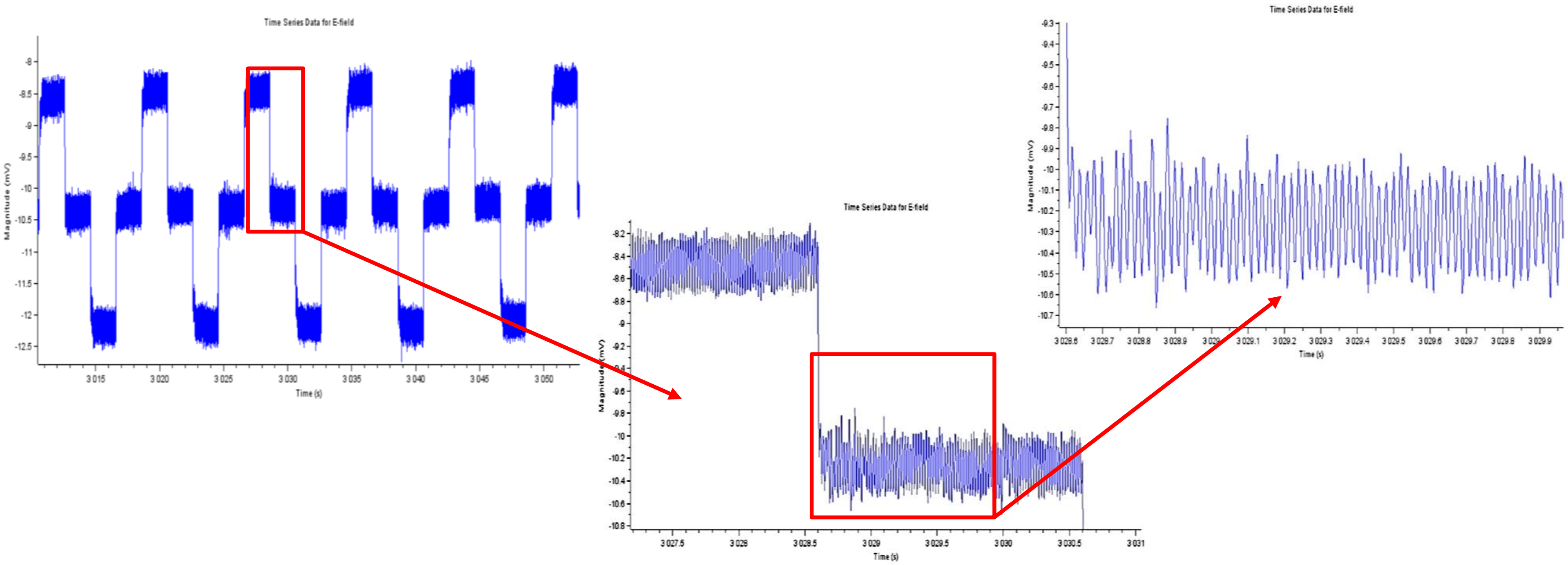
Inferrence and subtraction of telluric noise
Telluric Cancellation

Time Series record from an Induced Polarization using gDAS24 (Fs=256Hz) for a 200m dipole, extending over an interval of approximately 30min. The blue trace is the observed data, the green is the scalar Telluric Cancelled signal and the red trace is the full tensor Telluric Cancelled signal. In the main panel the mid to low period telluric noise clearly distorts the observed signal to an extent that would potentially bias the survey results. Although both telluric cancelled results significantly reduce the effect of apparent telluric signal, the full tensor application provides more accurate inference and removal.



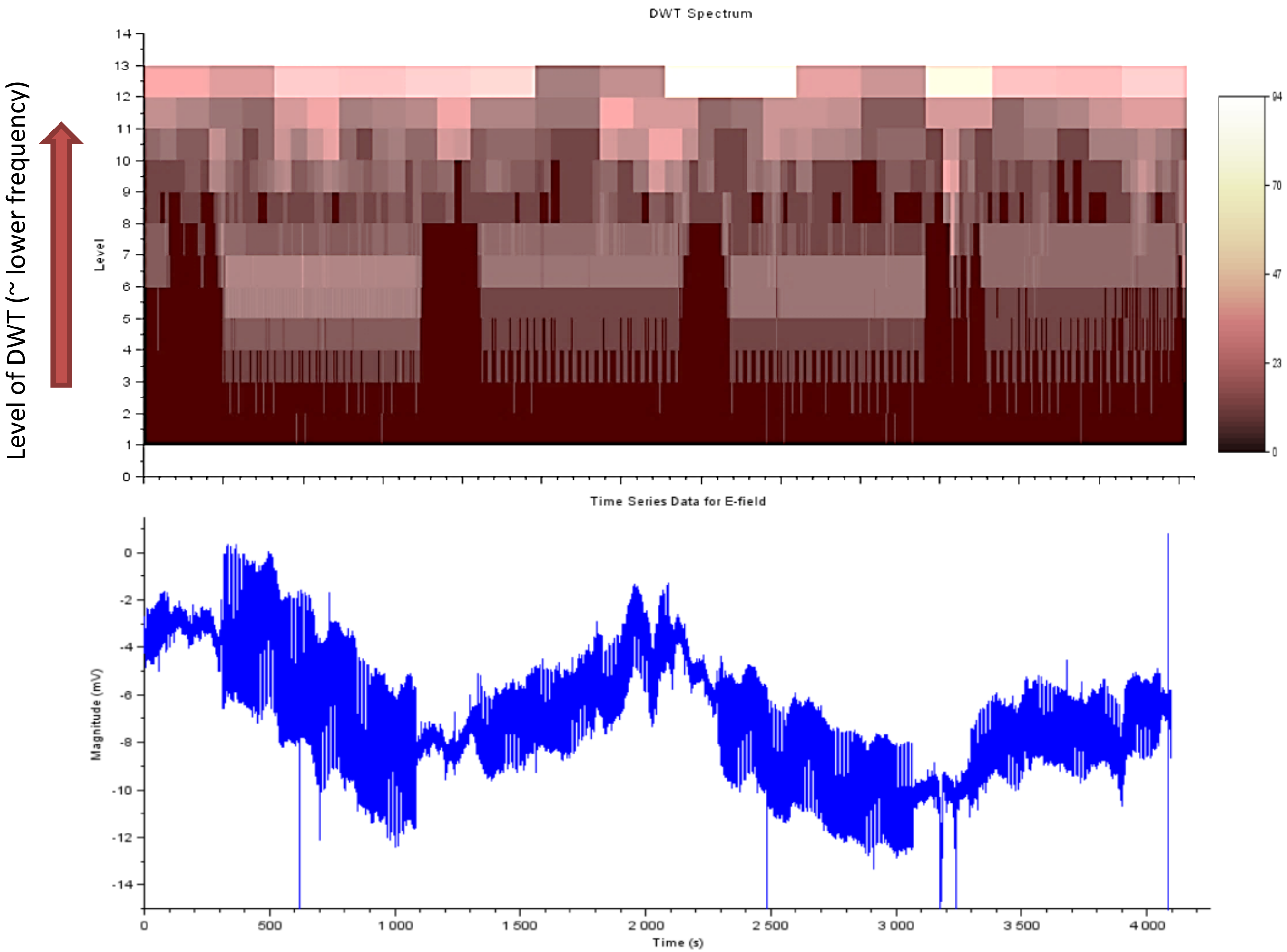
Following Telluric Cancellation
Time series filtering

- Aim is to remove remaining components of the (telluric cancelled) signal that may be considered as noise and as such are detrimental to the desired response
- FFT or Wavelet filtering
  - Which results best depends on the nature of the noise (And signal)
  - FFT – remove or taper frequency content below and/or above those of interest, say <0.01Hz and >20Hz usually using a Butterworth or similar filter design
  - Wavelet – remove low or upper levels of DWT to leave the required signal less contaminated by noise:

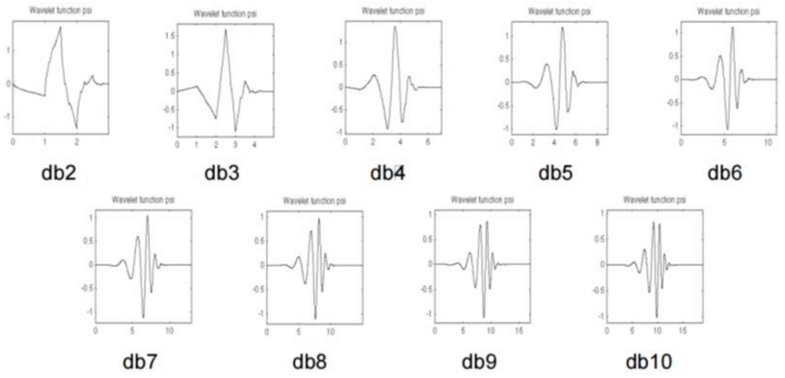


Time series filtering- Wavelets

Wavelets are often preferred given that they don't completely collapse the time series into the frequency domain, maintaining both such that the spectral content of the signal can be examined as a function of time. They also tend to require less preconditioning of the time series to work effectively.



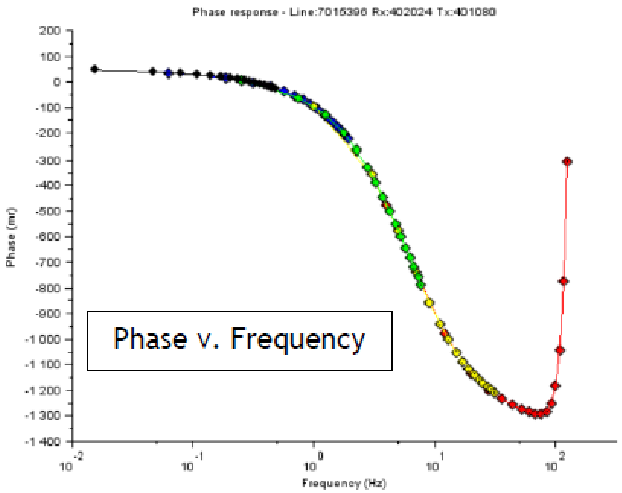
Typical is to use (soft) thresholded denoising with a "db8" for removal of high frequency noise and "db4" for low frequency variations (significantly below the frequency content of the measurable IP signal

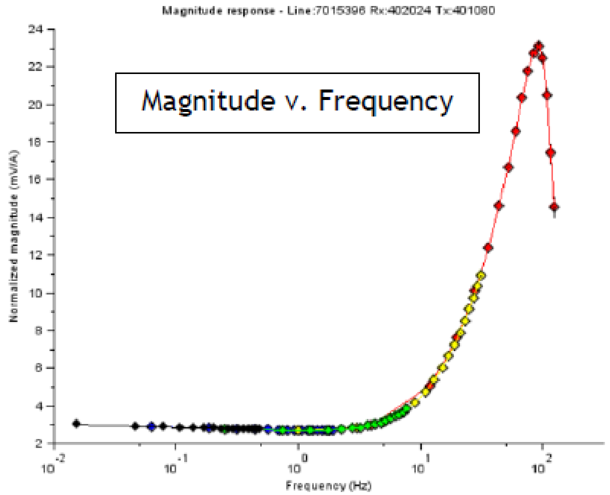


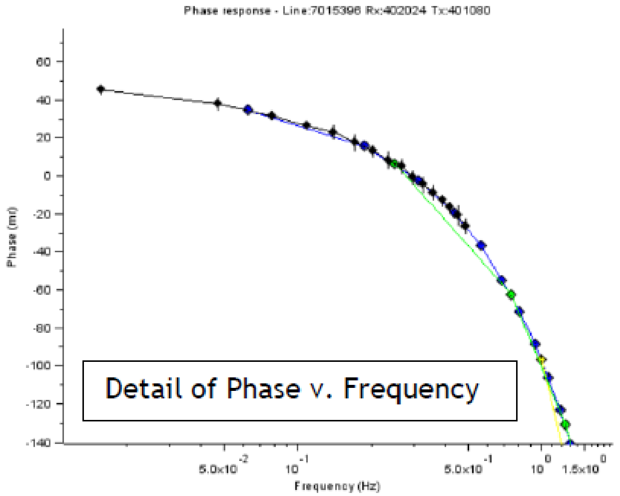
Selection of Processing Domain
Time or Frequency Domain

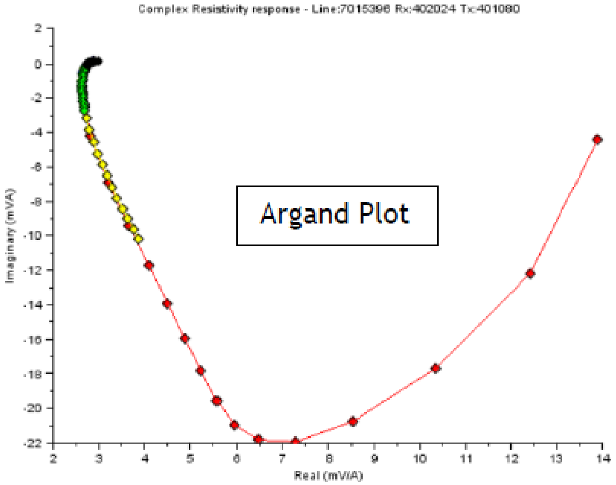
Option for both time and frequency domain processing is available

For Complex Resistivity (CR) the processing methodology has many analogies to the Magneto-Telluric (MT) processing stream, with the aim to obtain scalar or tensor transfer functions (magnitude and phase) based on the time series data from the current monitor (analogous to the H-field in MT) and the receiver dipole (analogous to the E-field in MT).







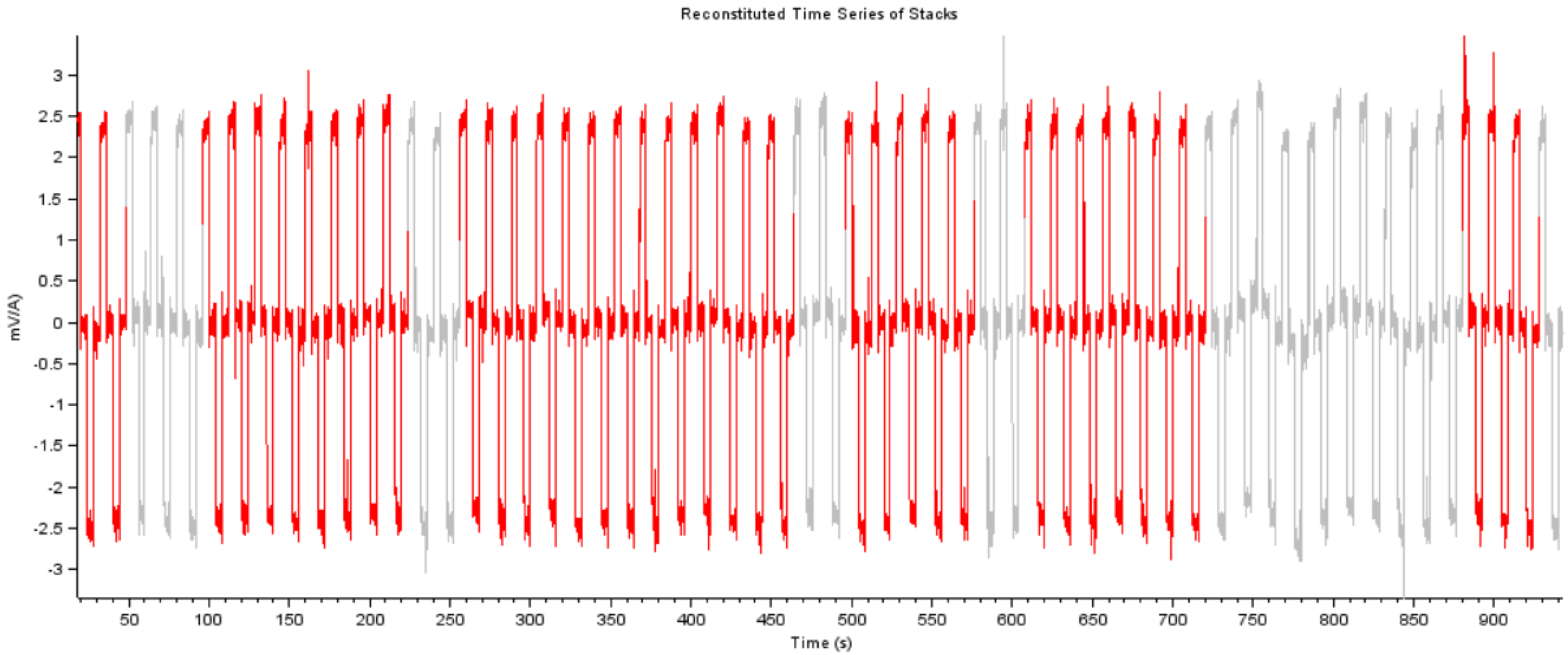


Processing in the Time Domain
Time series editting

The current monitor is analysed to identify each half cycle and its corresponding current intensity. Normalization of the E-field time series data for current is done on a cycle-by-cycle basis.

Prior to stacking, distorted looking data is identified and removed by evaluating the correlation of each cycle to an “ideal” waveform and then removing for example the 20% of the worst correlations. User selection/removal can be also be done graphically.

Manual selection / rejection is also posible in the software

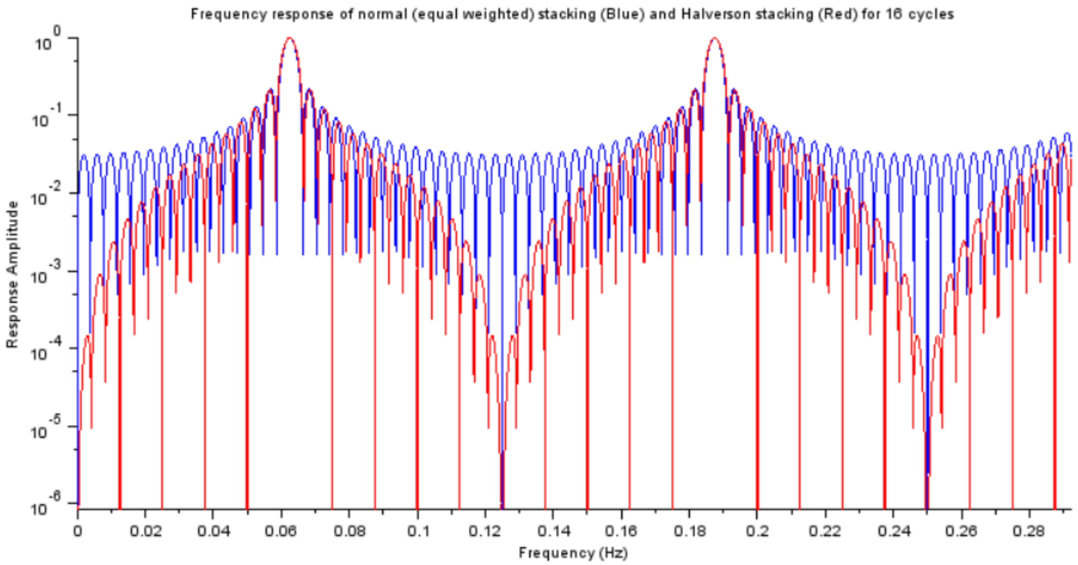


Processing in the Time Domain
Stacking and Averaging

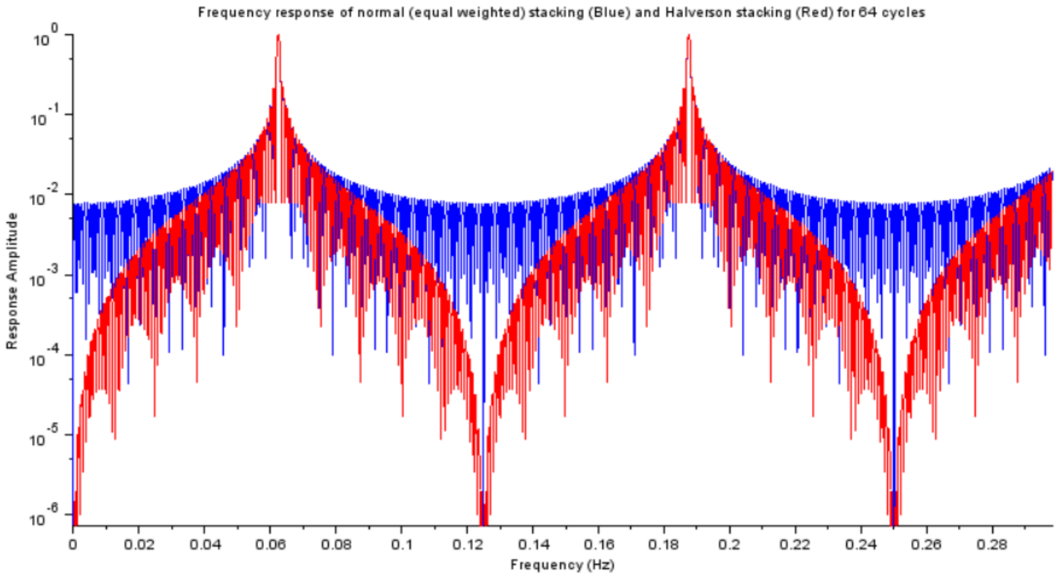
- Stacking and averaging is a filtering process, comprised of a moving average filter, with standard stacking based on a filter with coefficients of ½,-½ with each factor followed by enough zeros such that each is one half-cycle apart
  - "Halverson stacking" (factors of ¼, -½, ¼ for each of three consecutive half cycles) may provide significant benefits over standard stacking (with equal weightings on all half cycles) as seen in the filters frequency response with deeper stop bands (more effective noise rejection around even harmonics)

Standard (Blue) versus Halverson (Red) Stacking

Frequency response for 16 cycle stack



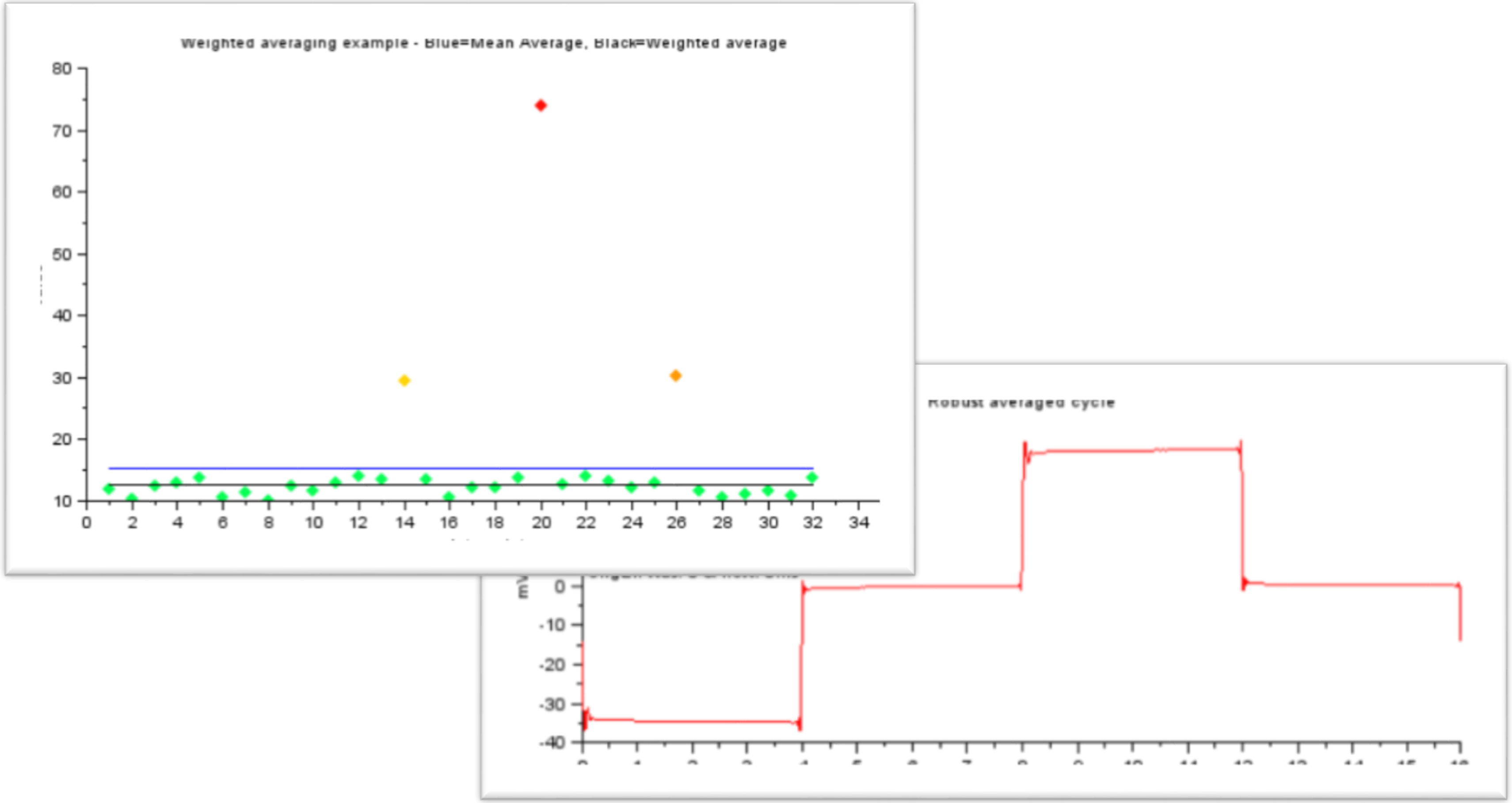
Frequency Response for 64 cycle stack



After J.Kingman,et.al., Terrigena Technical White Paper Number 20040328.0101(e) 2004

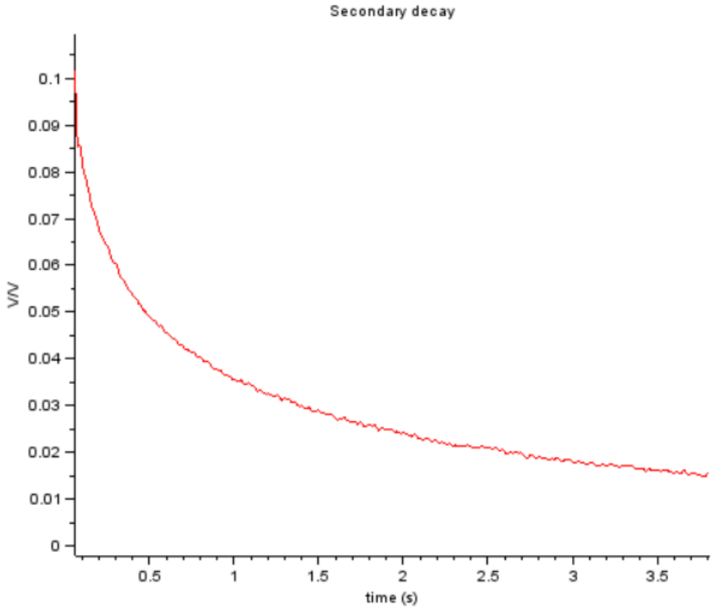
Processing in the Time Domain
Stacking and Averaging

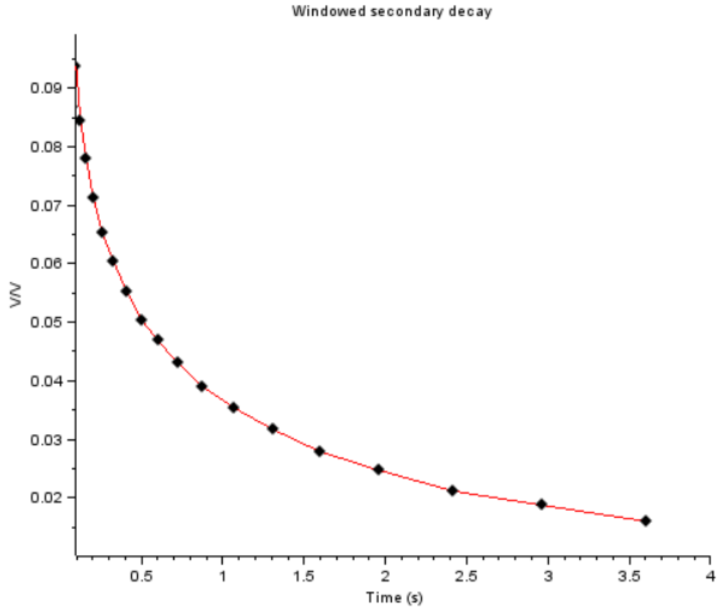
- For noisier data a robust average across the stack of cycles is used which weights data by the square of the inverse of the number of standard deviations of the sample value from the median value of all of the corresponding samples. The result is a robustly averaged full cycle. The two half cycles can then be stacked and averaged to provide the final decay.



Decay windowing

- Final windowing of decay and calculation of a chargeability
  - Windowing is used to reduce the secondary decay response to a more manageable number of data points from, for example, 1024 data points for a 4s off-time sampled at 256Hz to a user selected number of values (usually 20, more than adequate to describe the decay curve and sufficient for a visual evaluation of data quality).
  - Usually done on a semi-exponential basis so later window times with longer time constants are correspondingly longer
  - We usually use a 16s (0.0625Hz) 50% duty cycle waveform providing an off-time of 4s, and an integration interval of 1.5 to 3.5s (this can be moved even later still) to avoid or at least minimize inductive EM-coupling effects.





Data review and editting
PDIP, DDIP, BDIP, 3D,Vector, Tensor, Gradient, and other arrays

- Software written in Scilab
- Software to easily review results and permit user to mine back to time series record, edit, restack
- Option to compare non-telluric and telluric cancelled data
- Option to process as Complex Resistivity (CR)