Magnetotelluric data processing. From theory to practice.

Maxim Smirnov
Department of Earth Sciences,
Uppsala University, Sweden
E-mail: maxim.smirnov@geo.uu.se

Ivan Varentsov
IGEMI, Troitsk, Moscow
E-mail: igemi1@mail.transit.ru

Theoretical principles of MT data processing

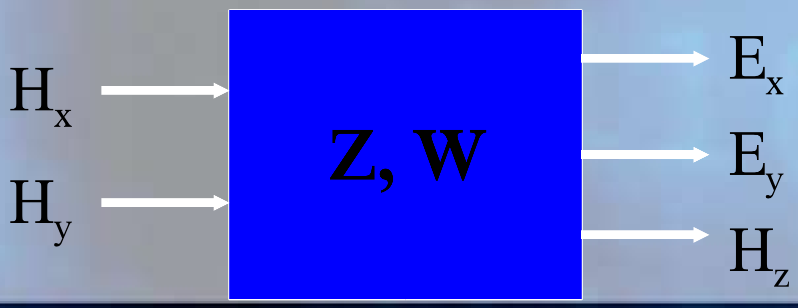
Maxim Smirnov
Department of Earth Sciences,
Uppsala University, Sweden
E-mail: maxim.smirnov@geo.uu.se

Summary

- Theoretical principles of MT data processing
  - Spectral analysis
  - Regression problem
- Synthetic tests
- Data examples

Magnetotellurics

- Input data are time series of natural electromagnetic field components measured at the Earth's surface.
- Magnetotelluric signals have a very wide frequency spectrum, caused at high frequencies by worldwide thunderstorm activity and at low frequencies by geomagnetic phenomena.



Earth as a linear system

Transfer functions to be estimated in frequency domain

MT transfer function: $\begin{pmatrix} E_x \\ E_y \end{pmatrix} = \begin{bmatrix} Z_{xx} & Z_{xy} \\ Z_{yx} & Z_{yy} \end{bmatrix} \begin{pmatrix} H_x \\ H_y \end{pmatrix}, \boldsymbol{E} = \boldsymbol{ZH}$

GDS transfer function $H_z = \begin{pmatrix} W_x & W_y \end{pmatrix} \begin{pmatrix} H_x \\ H_y \end{pmatrix}, H_z = \boldsymbol{W}^{\boldsymbol{T}} \boldsymbol{H}$

Horizontal magnetic transfer function

$$\begin{pmatrix} H_x \\ H_y \end{pmatrix} = \begin{bmatrix} M_{xx} & M_{xy} \\ M_{yx} & M_{yy} \end{bmatrix} \begin{pmatrix} H_{xr} \\ H_{yr} \end{pmatrix}, \boldsymbol{H} = \boldsymbol{M}\,\boldsymbol{H}_{\boldsymbol{r}}$$

Main steps in MT data processing

- Time series preprocessing
- Spectral analysis
- Statistical analysis
  - Least squares
  - Robustness
  - Robust estimators
- Post processing
  - Multi window averaging
  - Multi RR averaging

Time Series Preprocessing

- Prewhitening, detrend
  - AR-filter
  - Polynomial fit
  - **First differences filter:** $\bar{x}[i] = x[i] - x[i-1]$
- Spikes, steps removal
- Filtering, decimation
  - Recursive IR filtering
  - Power line harmonics filtering

Spectral Transformation

- Classical spectral estimations
- Window tapering
- Cascade Decimation
- Spectral Resolution

Classical spectral estimation

Periodogram method is based on Fast Fourier Transform of time series with following ensemble or frequency averaging.

- Daniell frequency averaging. Performed after FFT of a complete time series

$$S_{XY}(j)=\frac{1}{k+1}\sum_{m=-k/2}^{k/2} S^{n}_{XY}(j+m),\ j=\frac{k}{2}+1,\ldots,\frac{N}{2}-\frac{k}{2},$$

where $S^{n}_{XY}(j)=\frac{1}{T}X^{*}(j)Y(j)$ non-smooth auto and cross spectra.

Classical spectral estimation

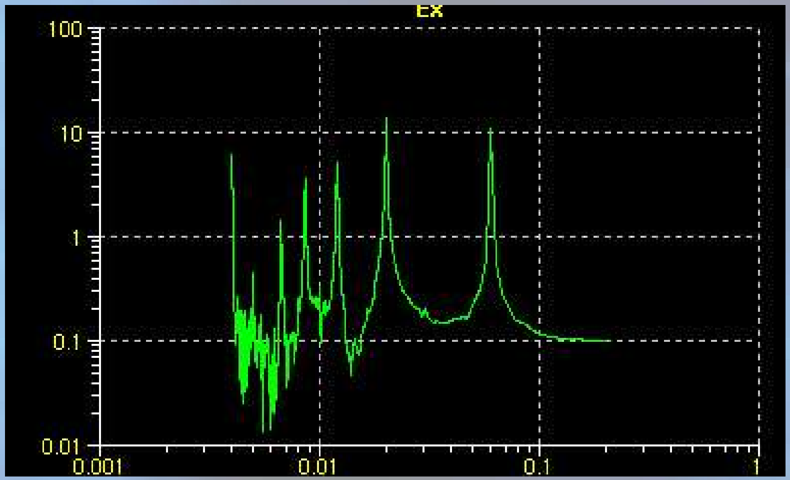
- Welch ensemble averaging. Time series are subdivided into overlapped segments. FFT is performed for each segment. Smooth spectra are estimated as follows:

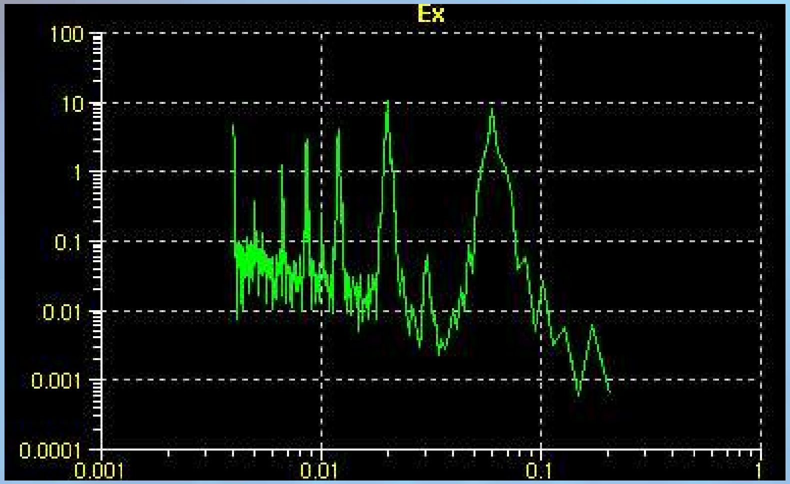
$$S_{XY}(j) = \frac{1}{M}\sum_{m=1}^{M} S^{n}_{XY}(j),$$ where M is a number of segments.

Window tapering

- To reduce spectral leakage and hence the bias, each segment is windowed by data taper. An example of widely used taper is Hanning window:

$$\bar{x}[i] = x[i]h[i]; h[i] = \frac{1}{2}\left(1 - \cos\frac{2\pi i}{N-1}\right)$$





AMT data sampled at 1000 Hz. Strong power line harmonics are seen.

Without tapering

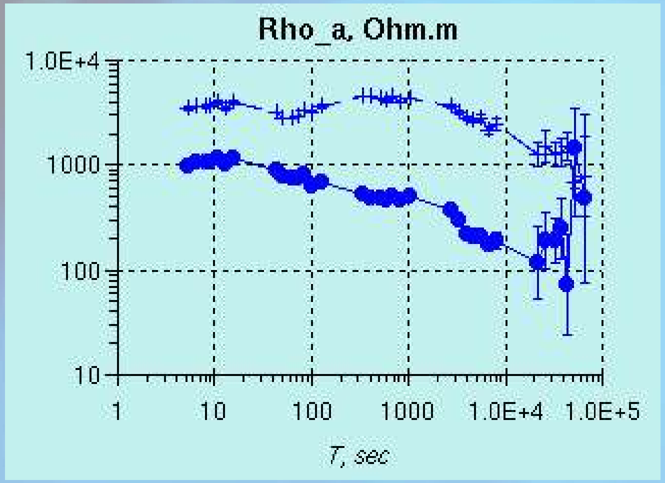
Hanning tapering window

The procedure

- Complete time series are subdivided into overlapping segments.
- The overlapping usually taken up to 50% when the segments might be still considered as independent.
- Each segment is subject to Fast Fourier Transform after detrend and tapering.
- Fourier harmonics or non-smooth cross- and auto-spectra are stored for further analysis.

Cascade decimation technique

The whole period range is covered by applying cascade decimation technique. It consists of several decimation steps.
Each decimation step involves antialiasing lowpass filtering of the time series with a recursive filter and then decimation.



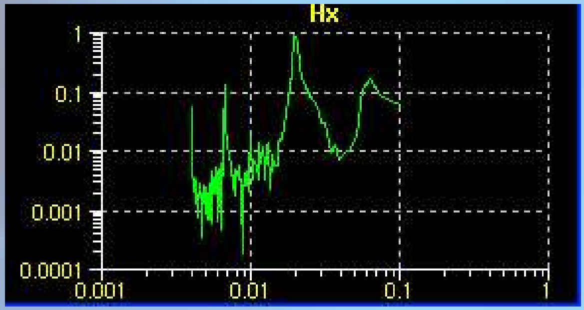
ρxy, ρyx for LMT site. Results from different decimation steps.

Advantages of the technique are

- fast computations
- covers a whole period range with the same FFT length

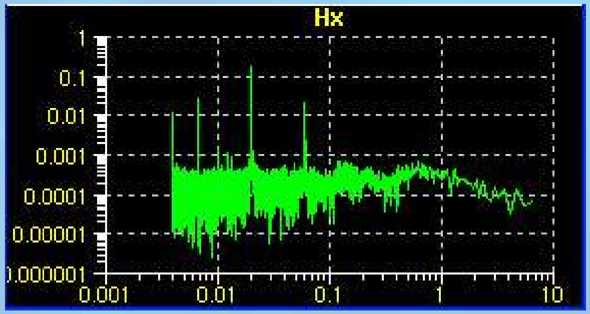
Spectral resolution

- The resolution of Fourier transform is roughly proportional to 1/T, where T is the length of data segment.



FFT length is 512.

- There are two opposite tendencies which influence the bias of the final estimates.
  - Short segments allow better sorting of noise contaminated parts of recording
  - But spectral estimates might be biased if strong harmonic signals are presented



FFT length is 65536

Length of Fourier Transform

- Usually short FFTs are efficient to process long period data and less sensitive to non-stationarity
- On the other hand long FFTs are required to process high frequency data (AMT range) to resolve and eliminate power harmonics.

Sorting algorithms

Most of the algorithms will work well unless a large part of the data are noisy. After spectral analysis only those segments which satisfy certain criterions will be used further in statistical procedures.

- Coherence sorting
- Other physical criterions
  - Size of Tipper components
  - Horizontal transfer function
  - Impedance phase

Coherence sorting

- Common coherence: $$\gamma_{xy}^2 = \frac{[\![S_{xy}]\!]^2}{S_{xx} S_{yy}}$$
- Partial coherence: $$\gamma_{x_1 y \cdot x_2}^2 = \frac{[\![S_{x_1 y \cdot x_2}]\!]^2}{S_{x_1 x_1 \cdot x_2} S_{yy \cdot x_2}}, \quad S_{xy \cdot z} = S_{xy} - S_{zy} S_{xz} / S_{zz}$$
- Multiple coherence: $$\gamma_{y:x}^2 = 1 - [1 - \gamma_{x_1 y}^2][1 - \gamma_{x_2 y : x_1}^2]$$

Statistical analysis

- Linear regression problem. Least squares solution.
- Gauss-Markov theorem and Least Squares estimate for regression problem
- Robustness. Breakdown point.
- Robust estimators.
- Error estimates
  - LS errors
  - Jackknife/Bootstrap methods

Linear regression problem

The system of MT equations poses a linear regression problem which, in general terms, can be written as:

$$y_i = \boldsymbol{x}_i^T \boldsymbol{\Theta} + e_i, i = 1, \ldots, M,$$

where y_i is the predicted (response variable) value from the i-th observation of a p-dimensional vector $\boldsymbol{x}_i$ (predictor),

e_i the i-th prediction error,

while $\boldsymbol{\Theta}$ represents the p-dimensional vector of unknown regression parameters to be estimated.

Linear regression problem

The complete system of equations for M observations:

$$\begin{pmatrix} y_1 \\ . \\ . \\ . \\ y_M \end{pmatrix} = \begin{bmatrix} x_{11} & \dots & x_{1p} \\ . & \dots & . \\ . & \dots & . \\ . & \dots & . \\ x_{M1} & \dots & x_{Mp} \end{bmatrix} \begin{pmatrix} \Theta_1 \\ . \\ . \\ . \\ \Theta_p \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ . \\ . \\ . \\ \varepsilon_M \end{pmatrix}$$

or concisely in matrix notation $\boldsymbol{y} = \boldsymbol{X}\,\boldsymbol{\Theta}$

Least Squares estimate for regression problem

The least squares (LS) solution T_M^{LS} of the thus formulated regression problem can be found by minimizing the Euclidian norm of residuals r_i :

$$\Gamma(\Theta) = \sum_{i=1}^{M} \left(\frac{(y_i - x_i^T \Theta)}{\boldsymbol{\sigma}} \right)^2 ,$$

where $\boldsymbol{\sigma}$ is a scaling parameter.
Solving the **normal equations** $\boldsymbol{X}^T \boldsymbol{y} = \boldsymbol{X}^T \boldsymbol{X} \boldsymbol{\Theta}$ with respect to $\boldsymbol{\Theta}$ yields $\boldsymbol{\Theta} = (\boldsymbol{X}^T \boldsymbol{X})^{-1} (\boldsymbol{X}^T \boldsymbol{y})$,

Quality of statistical estimator

- Unbiased: $$E[\hat{\Theta}]=\Theta$$
- Consistent: $$\forall \varepsilon>0, \lim_{M\to\infty} Prob[|\hat{\Theta}-\Theta|\geqslant\varepsilon]=0$$
- Efficient: $$E[(\hat{\Theta}-\Theta)^2]\leqslant E[(\tilde{\Theta}-\Theta)^2]$$
- Asymptotic Relative Efficiency (ARE)

$$ARE(\hat{\Theta},\tilde{\Theta})=\frac{A\,var(\hat{\Theta})}{A\,var(\tilde{\Theta})}$$

Gauss-Markov theorem and optimal properties of LS estimate

The **Gauss-Markov theorem** states that in a linear model in which the errors have expectation zero and are uncorrelated and have equal variances (homoscedastic),

- $E(e_i)=0,$
- $var(e_i)=\sigma^2<\infty,$
- $cov(e_i,e_j)=0, \quad i\neq j$

the best linear unbiased estimators of the coefficients are the least-squares estimators. The errors are **not** assumed to be independent and normally distributed, nor are they assumed to be identically distributed.

Why robust statistics

- Linear model is not completely adequate
  - Source field effect => the relations between field components are not linear anymore.
- Frequently large residuals,
  - Cultural noise
  - High amplitude disturbances
- Any of these problems may easily destroy the optimal properties of LS

Robustness

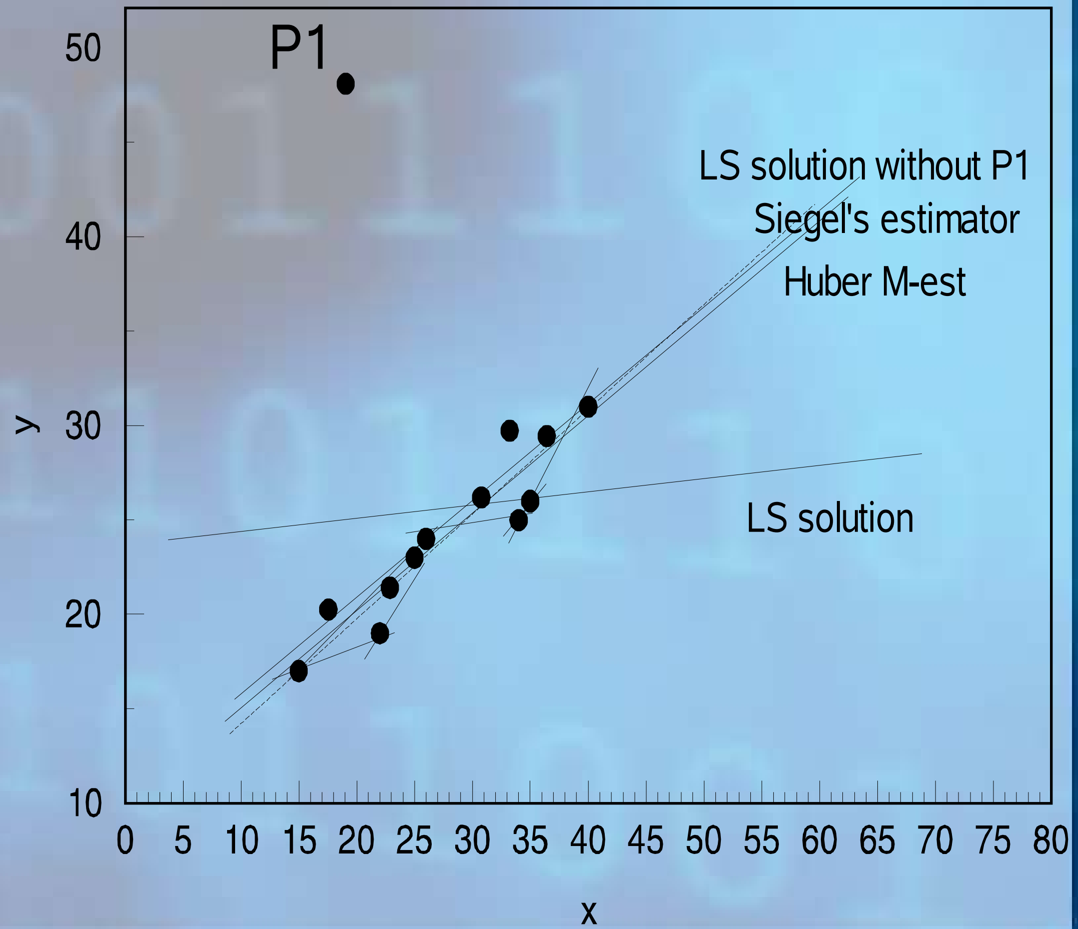
The basic measure of the robustness of an estimator is its **breakdown point** ε^*, that is, the fraction (up to 50%) of outlying data points that can corrupt the estimator.

In other words, the breakdown point is the smallest proportion of contaminated observations that can carry the estimator beyond all bounds. This contamination is not restricted to the ouliers in response variables, but that predictors may also contain ouliers.

LS solution has a breakdown point equal to zero, which means that even a few outliers in original data may seriously corrupt the final result.

Robustness

Effect of a gross error P1 in the data, for simple linear regression problem. LS solution is completely destroyed by one outlier. Robust estimates are much less influenced by this outlier.



Huber's regression M-estimate

Huber suggested to minimize the non-quadratic loss function:

$$\Gamma(\Theta)=\sum_{i=1}^{M}\rho\left(\frac{(y_i-x_i^T\Theta)}{\sigma}\right)$$

or after taking derivatives $\sum_{i=1}^{M}\psi\left(\frac{(y_i-x_i^T\Theta)}{\sigma}\right)x_{ik}=0$

where $\psi(x)$ is referred to as the influence function.

Huber's M-estimator can be derived by putting

$$\psi(x)=\begin{cases}-c, & x\leq -c\\ x, & -c<x<c\\ c, & x\geq c\end{cases}$$

Huber's M-estimate

It leads to weighted LS problem $\sum_{i=1}^{M} w_i \xi_i x_{ik} = 0$ defined by the weights $w_i = min\{1, c/|\xi_i|\}$, $\xi_i = \dfrac{r_i}{\sigma}$, where r_i is i -th residual, c is a positive constant, σ is a scale estimate, which itself has to be robust.

Usual choice for σ is median of absolute deviations from median (MAD):

$$S_{MAD} = \underset{k}{med}\{|r_k - r_{med}|\}$$

It is shown that, in practice, the breakdown point of such an estimator does not exceed 30%. However, M-estimators are much less sensitive to outliers than LS estimator.

Hat matrix

LS solution is defined as follow $\boldsymbol{\Theta}=(\boldsymbol{X}^T\boldsymbol{X})^{-1}(\boldsymbol{X}^T\boldsymbol{y})$. The fitted values or LS estimates $\hat{y}_i$ of the observed values y_i are given by

$\hat{\boldsymbol{y}}=\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}(\boldsymbol{X}^T\boldsymbol{y})=\boldsymbol{H}\boldsymbol{y}$, where $\boldsymbol{H}=\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T$

The ***H*** is often called “hat matrix”. Its a symmetric projection matrix, diagonal elements of which satisfy

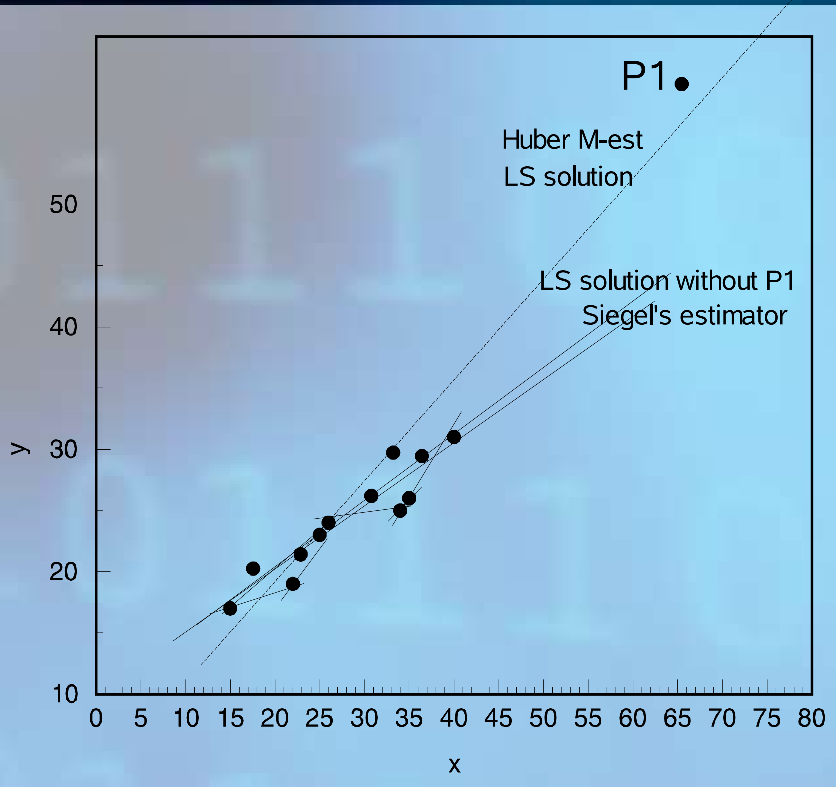
$0\leq h_{ii}\leq 1$; $tr(H)=p$; $ave\,h_{ii}=p/N$;

Leverage points

The *i*-th residual can be written

$$r_i = y_i - \hat{y}_i = (1 - h_{ii})\, y_i - \sum_{k \neq i} h_{ik}\, y_k$$

hence if h_{ii} is close to 1 (*i-th* observation is fitted perfectly), gross error in y_i will not necessarily show up in r_i . Points with large h_{ii} are leverage points.



Data example of simple linear regression with one leverage point P1, which is also an outlier. Its not possible to identify this point by the residual.

Bounded influence estimator

To reduce the influence of leverage points the bounded influence estimators are suggested. The following estimator bounds the influence of leverage points independently of corresponding residual,

$$\sum_{i=1}^{M} \psi\left(\frac{(y_i - x_i^T \Theta)}{\sigma}\right) v_i x_{ik} = 0 ,$$

where the weights v_i are determined based on the hat matrix.

Estimator having a high breakdown point

Siegel's repeated median estimator posses the highest possible breakdown point equal to $\varepsilon^* = 50\,\%$, which is expressed as follows:

$$T_n^{(j)} = \underset{i_1}{med}\{\ldots\{\underset{i_{p-1}}{med}\{\underset{i_p}{med}\{\Theta^{(j)}(i_1,\ldots,i_p)\}\}\}\ldots\},$$

where $\Theta^{(j)}(i_1,\ldots,i_p)$ is the j-th component of the unknown p-dimensional vector parameter, unequivocally determined by any p observations and $i=1,\ldots,n$ is the index of observation.

Repeated medians calculations

It is helpful to consider the simple linear regression model $y_i = \Theta_1 + \Theta_2 x_i + e_i$, to explain the estimator in details.

For each point (x_i, y_i), let's denote the median Θ_{2_i} of the $n-1$ slopes of the lines passing through this point and each other point of the set. The repeated median slope estimate Θ_2^* is defined to be the median of the multi set $\{\Theta_{2_i}\}$: $\Theta_2^* = \underset{i}{med}\, \underset{j \neq i}{med} \dfrac{y_i - y_j}{x_i - x_j}$.

The intercept Θ_1 can be estimated then either separately from Θ_2, as $\Theta_1^* = \underset{i}{med}\, \underset{j \neq i}{med} \dfrac{y_i x_j - y_j x_i}{x_j - x_i}$ or else hierarchically, as $\Theta_1^* = \underset{i}{med}\{y_i - \Theta_1^* x_i\}$

LS bias due to uncorrelated noise in magnetic components

The standard impedance LS estimate is biased due to noise in input channels which causes an over-estimate of the magnitude of the magnetic field auto-spectra.

$$Z_{xy} = (S_{ExHy})(S_{HyHy})^{-1} = (S_{Ex\tilde{H}y} + S_{ExN})(S_{\tilde{H}y\tilde{H}y} + S_{NN})^{-1},$$

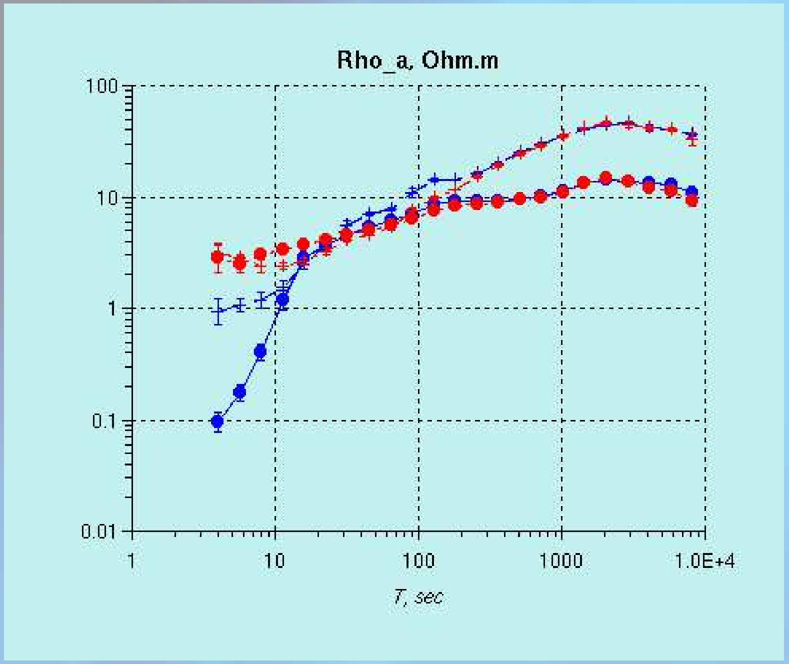
where $H_y = \tilde{H}_y + N, cov(\tilde{H}_y, N) = 0, cov(E_x, N) = 0$

Remote reference (RR) technique

To avoid the bias the RR method was developed:

$$Z = (E\,H_r^H)(H\,H_r^H)^{-1}$$

It also reduces the biases due to local correlated noise.



Rhoxy, Rhoyx; blue – SS estimate, red - RR

Regression in MT case

The regression problem for the first row of impedance tensor:

$$\begin{pmatrix} E_{x1} \\ . \\ . \\ . \\ E_{xN} \end{pmatrix} = \begin{bmatrix} H_{x1} & H_{y1} \\ . & . \\ . & . \\ . & . \\ H_{xN} & H_{yN} \end{bmatrix} \begin{pmatrix} Z_{xx} \\ Z_{xy} \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ . \\ . \\ . \\ \varepsilon_N \end{pmatrix}$$

Least squres (LS) solution: $\boldsymbol{Z} = (\boldsymbol{H}^{\boldsymbol{H}} \boldsymbol{H})^{-1} (\boldsymbol{H}^{\boldsymbol{H}} \boldsymbol{E})$

Remote reference form: $\boldsymbol{Z}_{\boldsymbol{r}} = (\boldsymbol{H}_{\boldsymbol{r}}^{\boldsymbol{H}} \boldsymbol{H})^{-1} (\boldsymbol{H}_{\boldsymbol{r}}^{\boldsymbol{H}} \boldsymbol{E})$

Remote reference weighted form with weights defined, for instance, by Huber's estimator: $\boldsymbol{Z}_{\boldsymbol{r}} = (\boldsymbol{H}_{\boldsymbol{r}}^{\boldsymbol{H}} \boldsymbol{W} \boldsymbol{H})^{-1} (\boldsymbol{H}_{\boldsymbol{r}}^{\boldsymbol{H}} \boldsymbol{W} \boldsymbol{E})$

LS errors estimation

Standard LS covariance matrix is

$$cov(\Theta) = \sigma^2 (\boldsymbol{X}^T \boldsymbol{X})^{-1}$$

where unbiased maximum-likelihood estimate for σ^2 is

$$s^2 = \|\boldsymbol{r}\|^2 / (N - p)$$

confidence limits are calculated as

$$\Delta_{P,n}(\Theta_k) = t_{(1-P)/2}(N - p)\, s\, (\boldsymbol{X}^T \boldsymbol{X})_{kk}^{-1/2},\ k = 1, \ldots, p$$

where t is Students t-distribution with *N-p* degrees of freedom, P probability level.

Resampling techniques to estimate errors

Jackknife:

- The jackknife (or leave-one-out) resampling technique aims at providing a computational procedure to estimate the variance and the bias of a generic estimator.
- The technique is based on removing samples from the available data set and recalculating the estimate.

Bootstrap:

- Bootstrap is a data-based simulation method for statistical inference.
- The bootstrap estimate of standard errors requires no theoretical calculations and is available no matter how mathematically complicated the estimator may be.

MT data processing algorithm by Smirnov

- Preprocessing
  - Detrend/prewitening with **first differences filter**
- Spectral analysis
  - Cascade decimation
  - Set of windows of different length (64-524288)
  - Segment overlapping is dependent on decimation step
  - Coherence frequency/segment sorting with automatically calculated coherence threshold
  - Hanning tapering window
- Statistical procedure (linear regression problem)
  - **Siegel's estimator.** Limits the influence of outliers in response variables (E field) as well as in factors space (predictors H field)
- **MultiRR Multi window averaging**

Multi window averaging

- After final processing the set of estimations for different Fourier length (for which different processing parameters were used) is available.
- At this stage all available partial estimates for particular site are averaged together (including different frequency bands, possibly from different instruments) to produce the final estimation.
- Averaging is performed with robust M-estimator and following bootstrap confidence limits.

Multi RR processing

- The use of several simultanious remote reference sites gives better control over distortions caused by the local noise and noise at remote site.
- Simple multi RR approach is based on robust averaging of all available RR estimations for selected site. The approach was thoroughly tested and applied in BEAR project.

Conclusions

- The most common approach to derive spectral estimates is based on FFT
- Robust statistics is a standard tool in MT data processing
- Remote reference technique significantly improves data quality especially in highly populated areas.
- Synchronous multi site averaging is the way to estimate additional transfer functions as well as to improve the reliability of the final estimates.

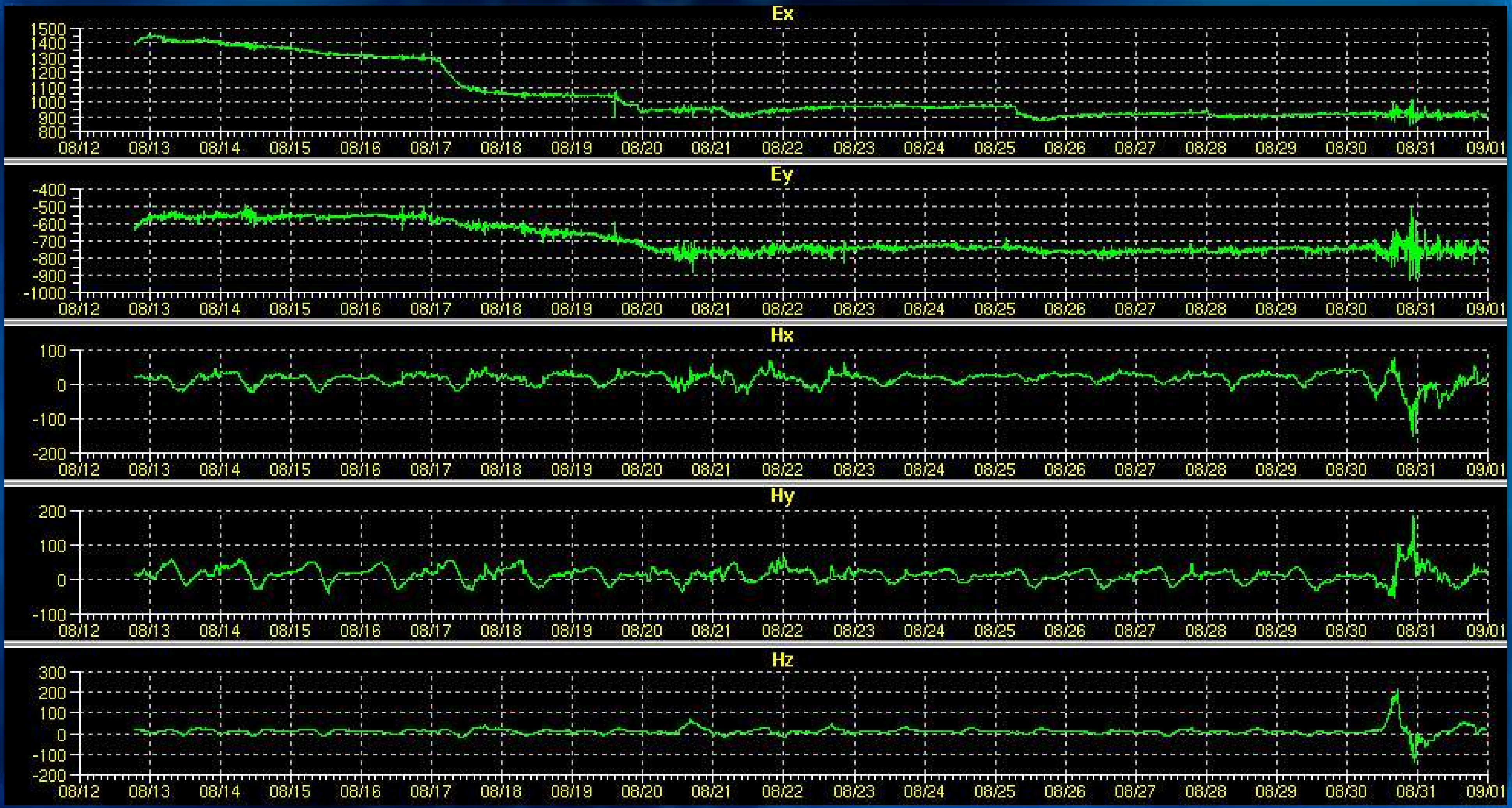
Data Examples

- Detailed example of data processing for LMT site
- COMDAT synthetic data set
  - Data set description
  - Tests example

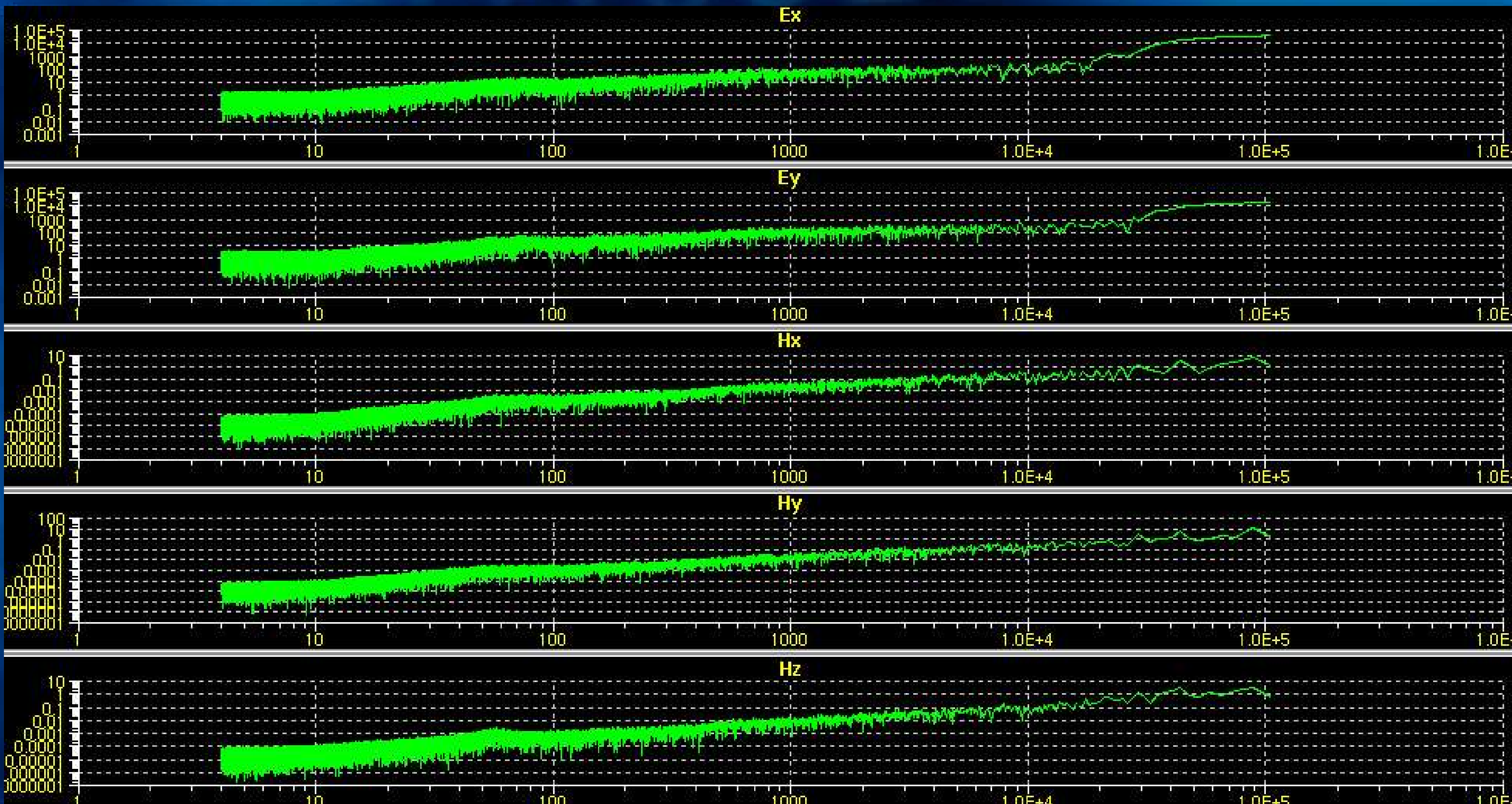
Data example

- LMT data processing. 1 sec sample rate. 19 days long. 8 sites were recorded simultaneously.
- AMT data processing. 2 simultaneous sites. Dual sample rate. 20 Hz continuos 1 day recording. 1000 Hz sample rate – 2h recording during night time.

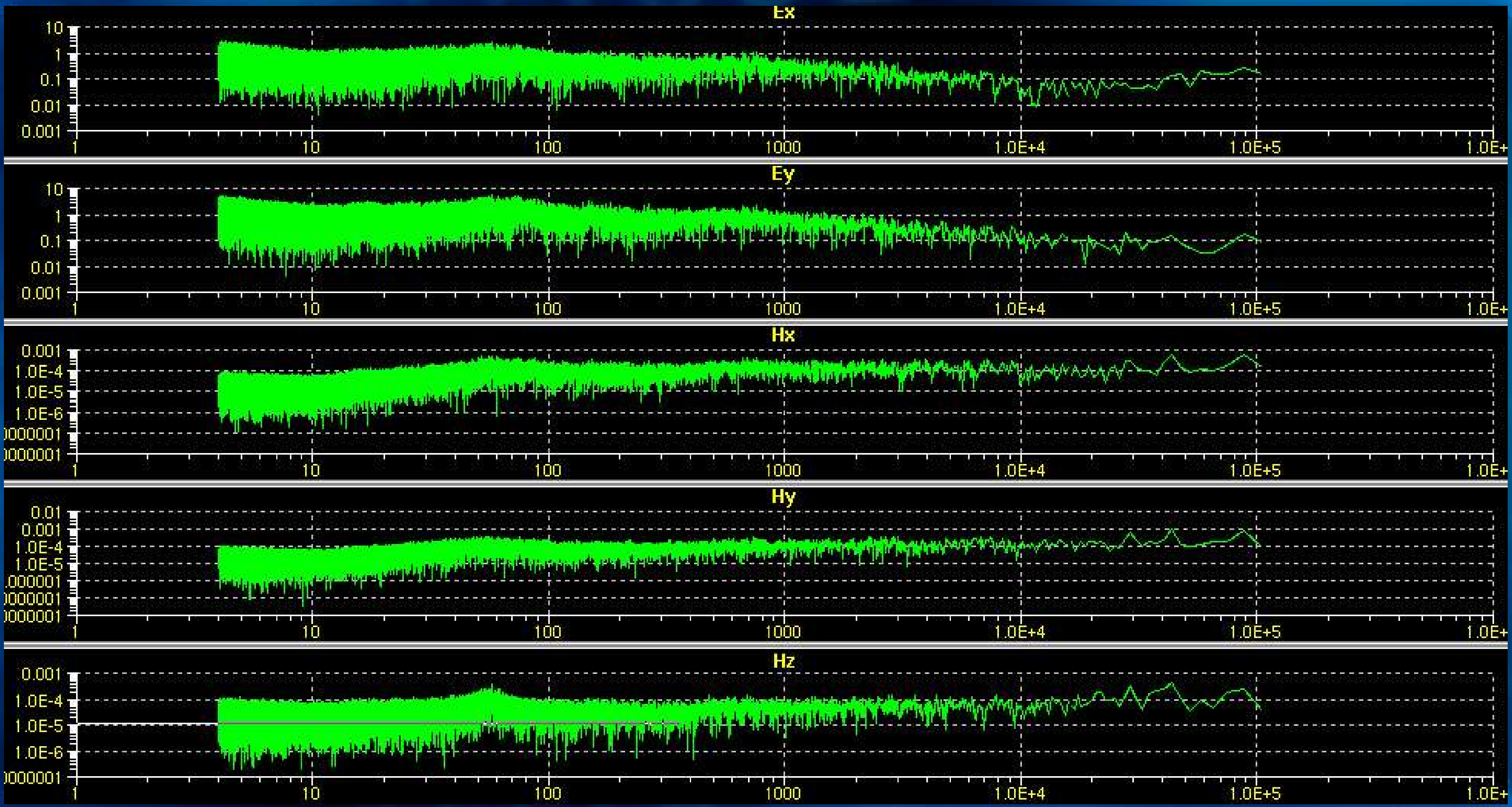
Raw LMT data


Ex
1500
1400
1300
1200
1100
1000
900
800
Ey
-400
-500
-600
-700
-800
-900
-1000
Hx
100
0
-100
-200
Hy
200
100
0
-100
Hz
300
200
100
0
-100
-200
08/12 08/13 08/14 08/15 08/16 08/17 08/18 08/19 08/20 08/21 08/22 08/23 08/24 08/25 08/26 08/27 08/28 08/29 08/30 08/31 09/01

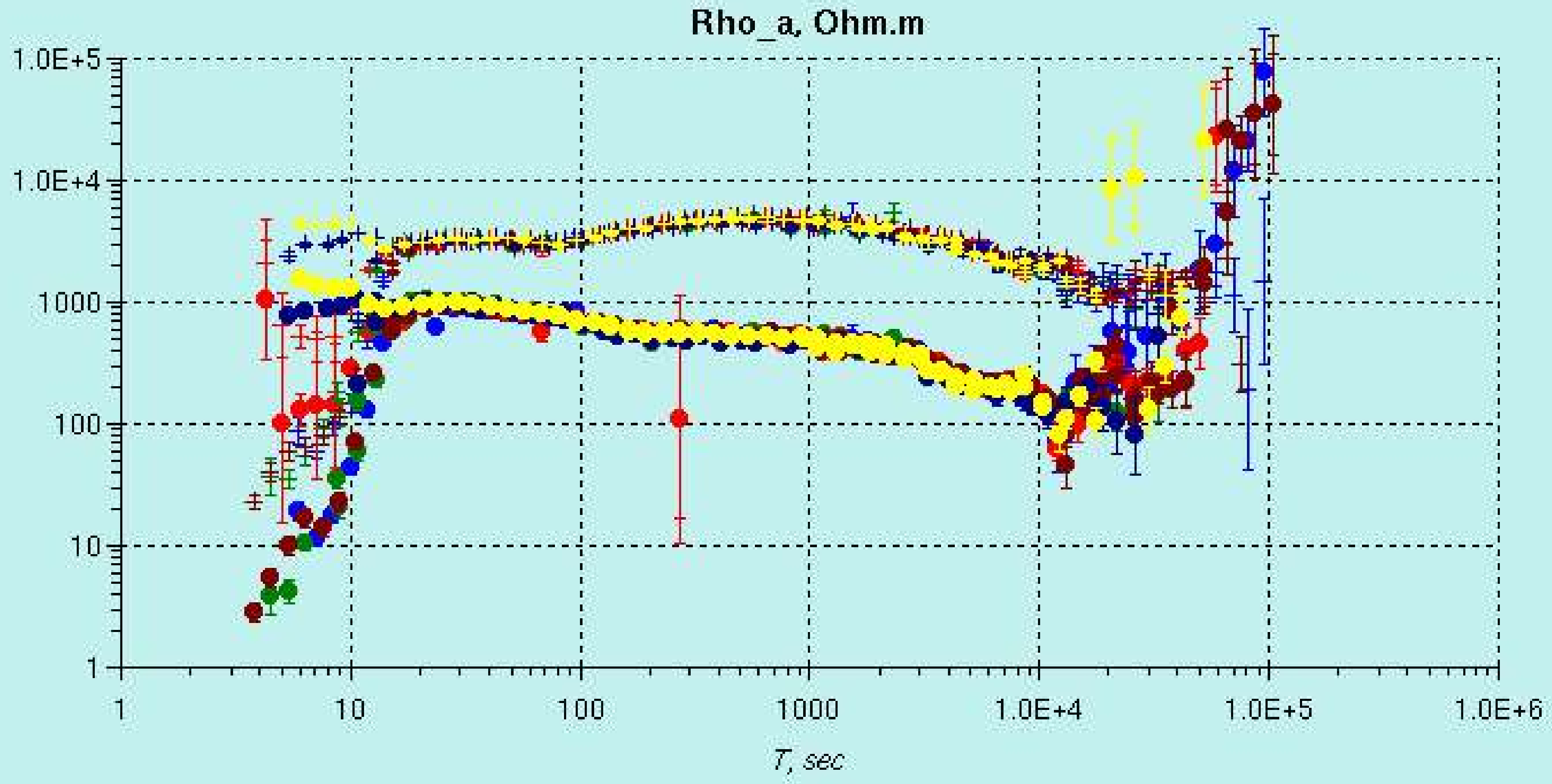
Raw LMT spectra



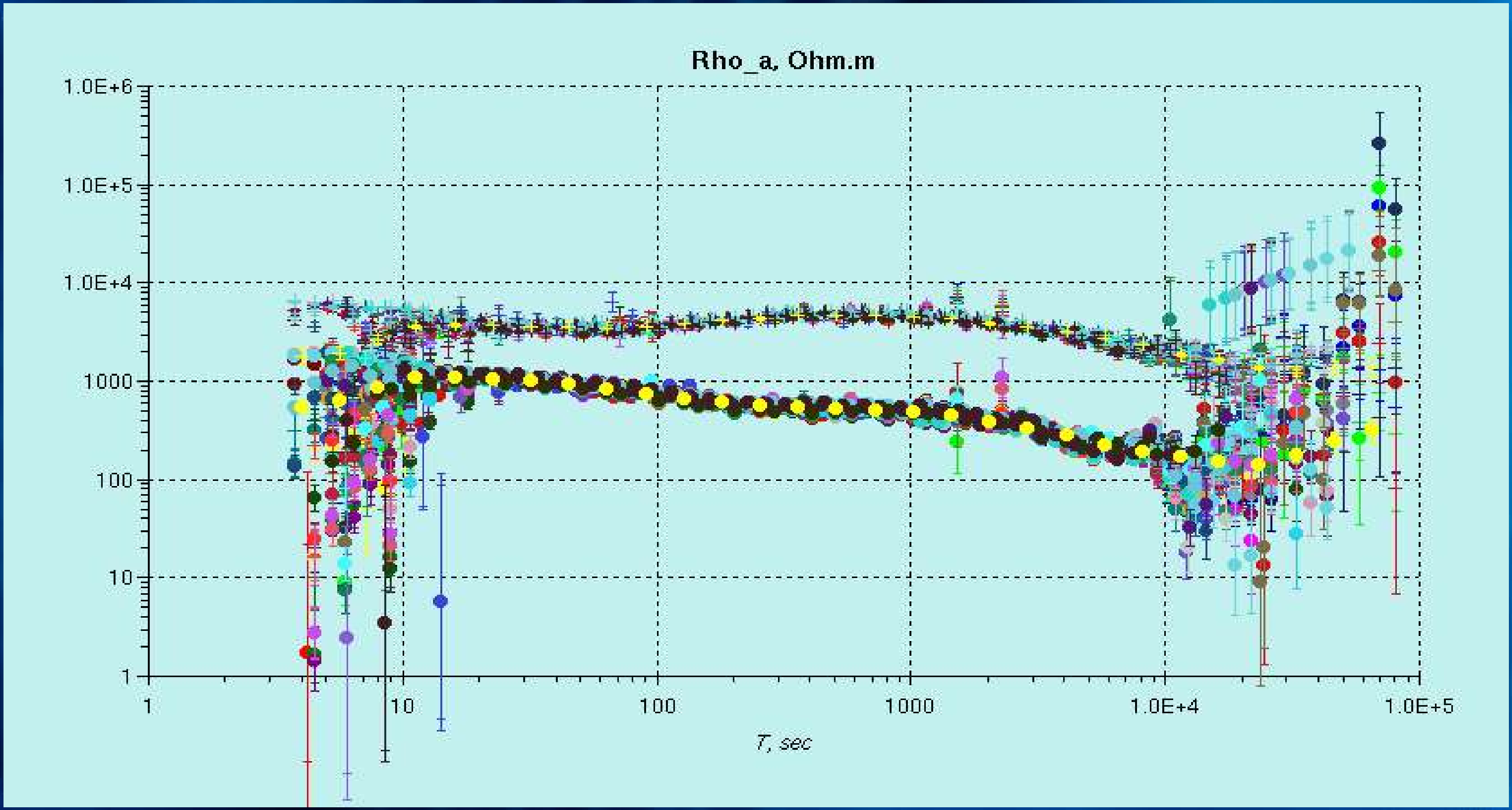
Raw LMT spectra with first deferences filter applied. “Prewhitening/detrend”



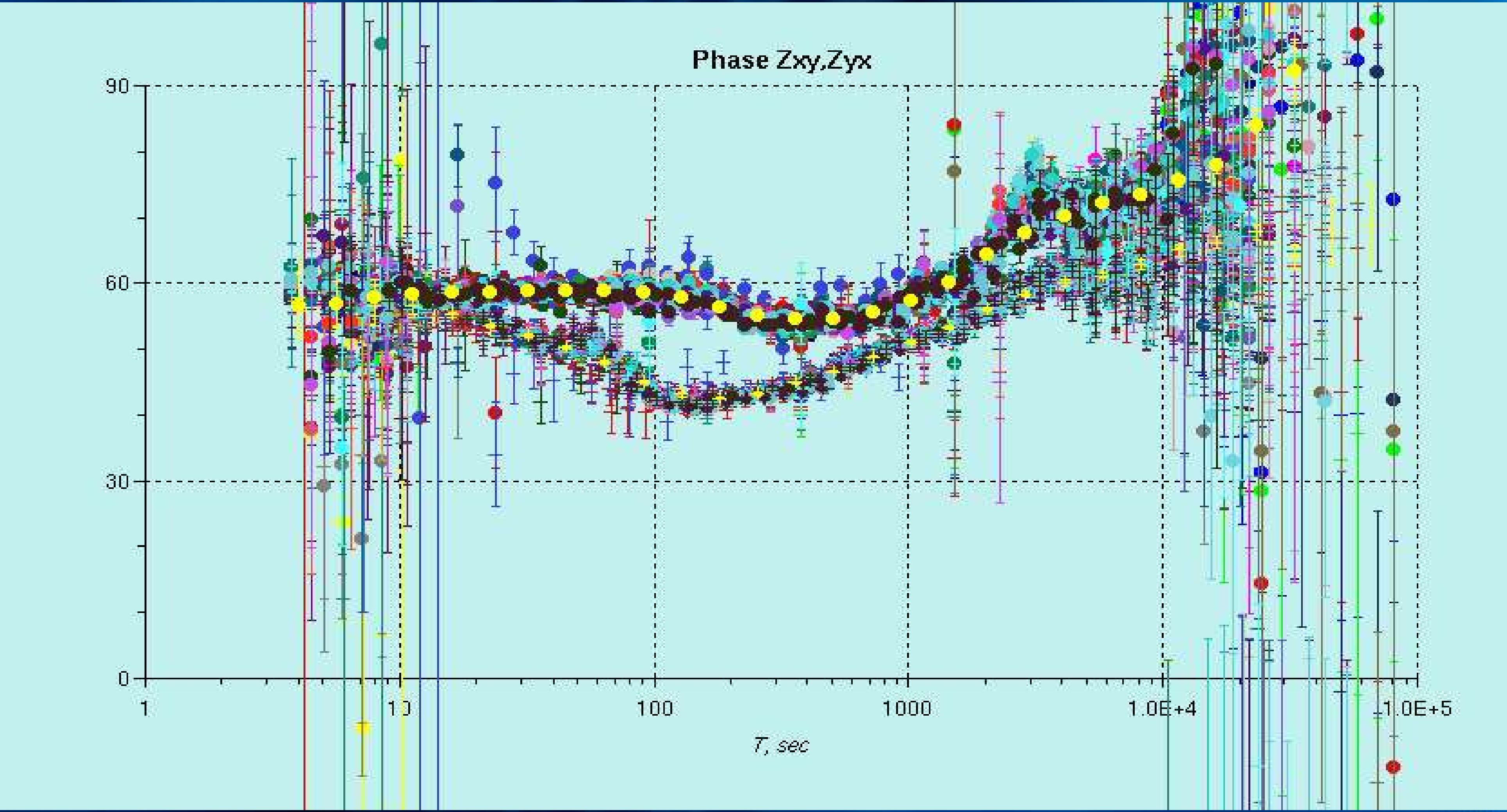
The set of SS estimates with different FFT length. Rhoxy, Rhoyx. Colors – for different Fourier length.



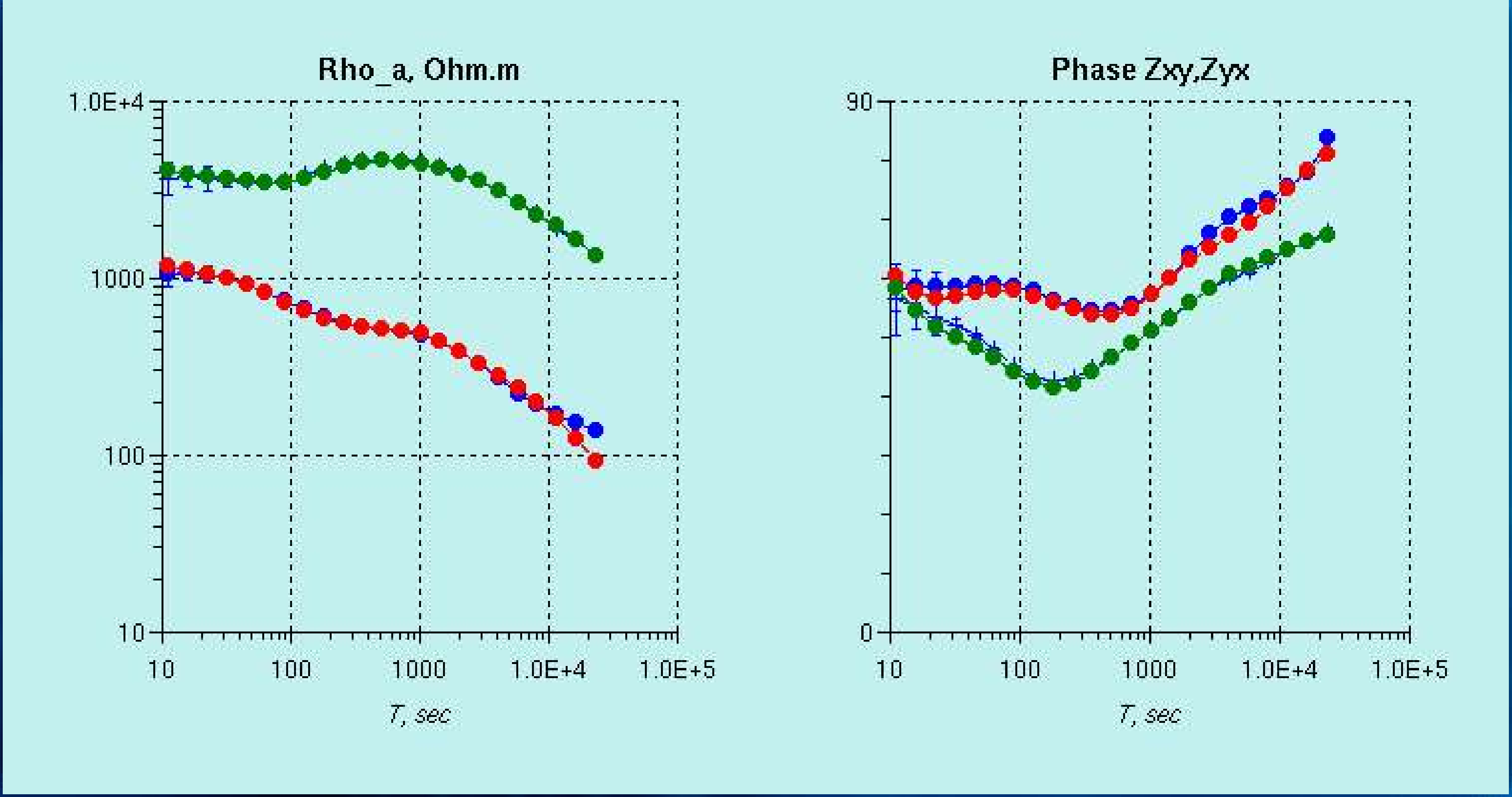
The same site. A complete set of all RR estimates obtained with 7 remote sites and 6 FFT lengths (64 - 65536)



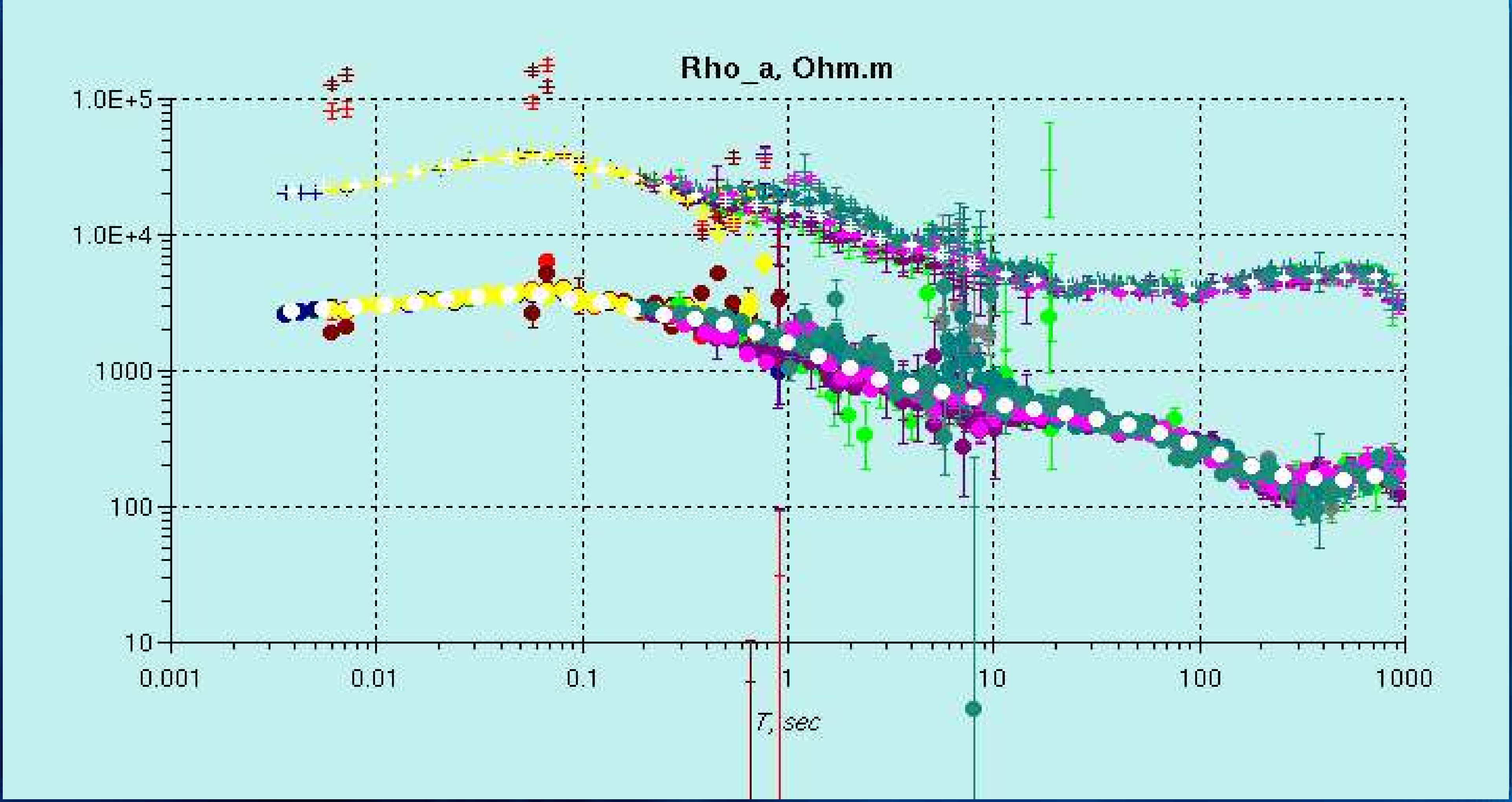
The set of phase estimates for the same site.



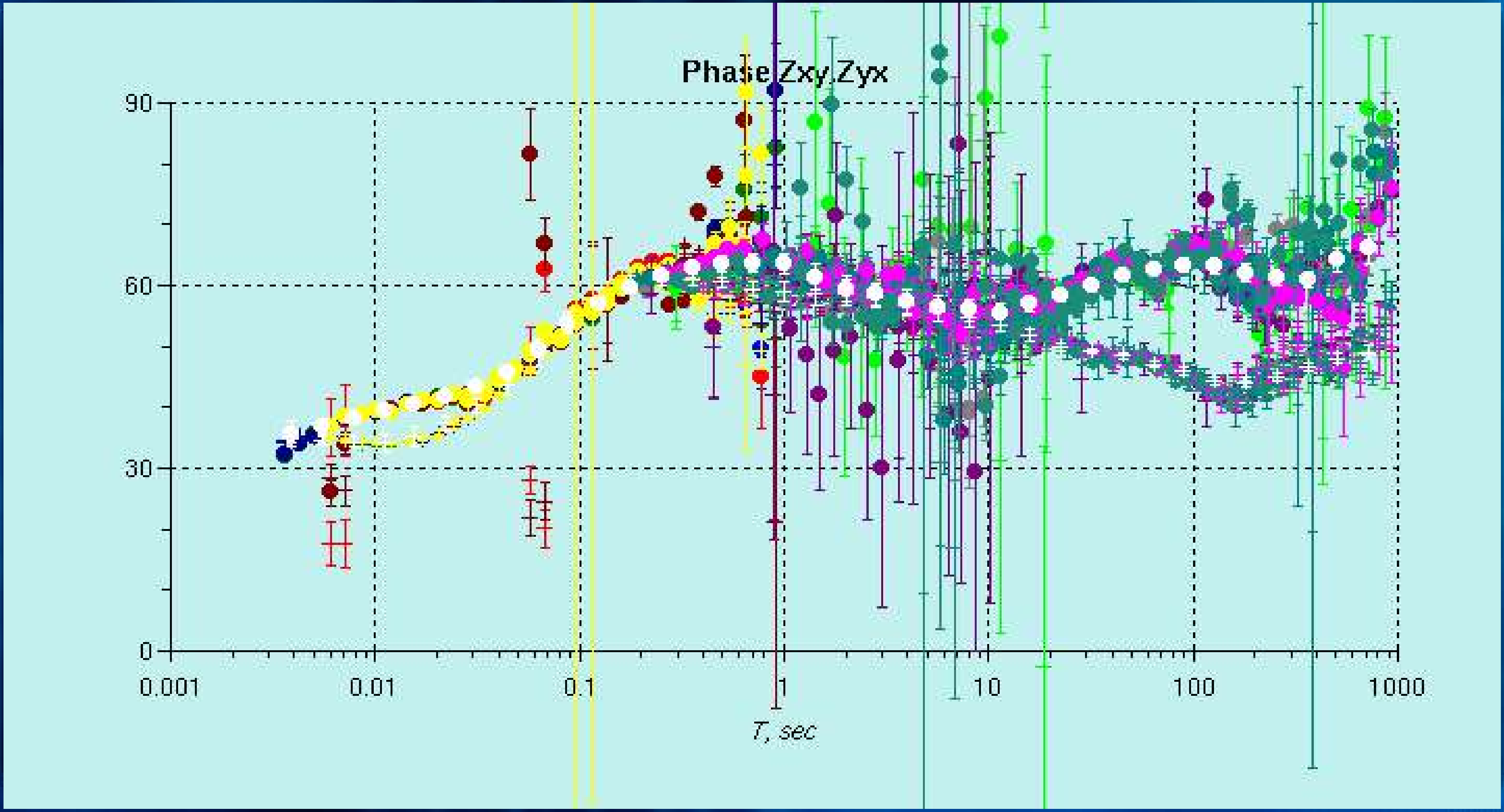
1D check for consistency (dispersion relations) between Rho and phase. Blue – the final average. Red and green are 1D model fit.



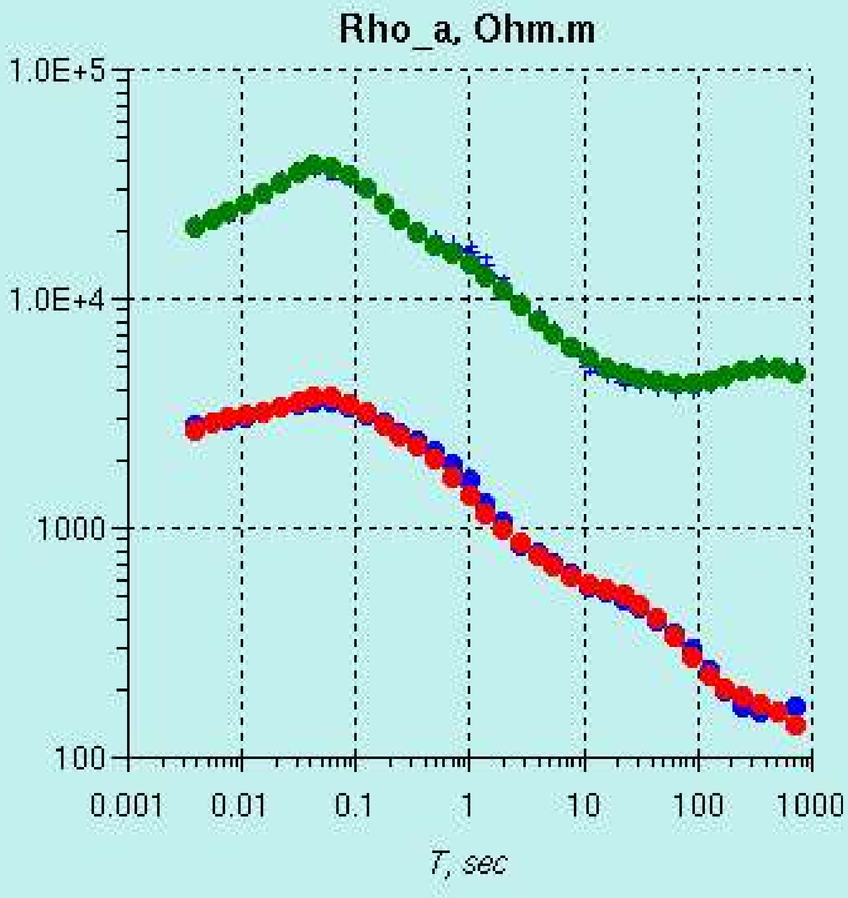
Complete set of RR estimations for AMT site (0.001- 1000 sec). 6 windows (64-262000) and one RR site.

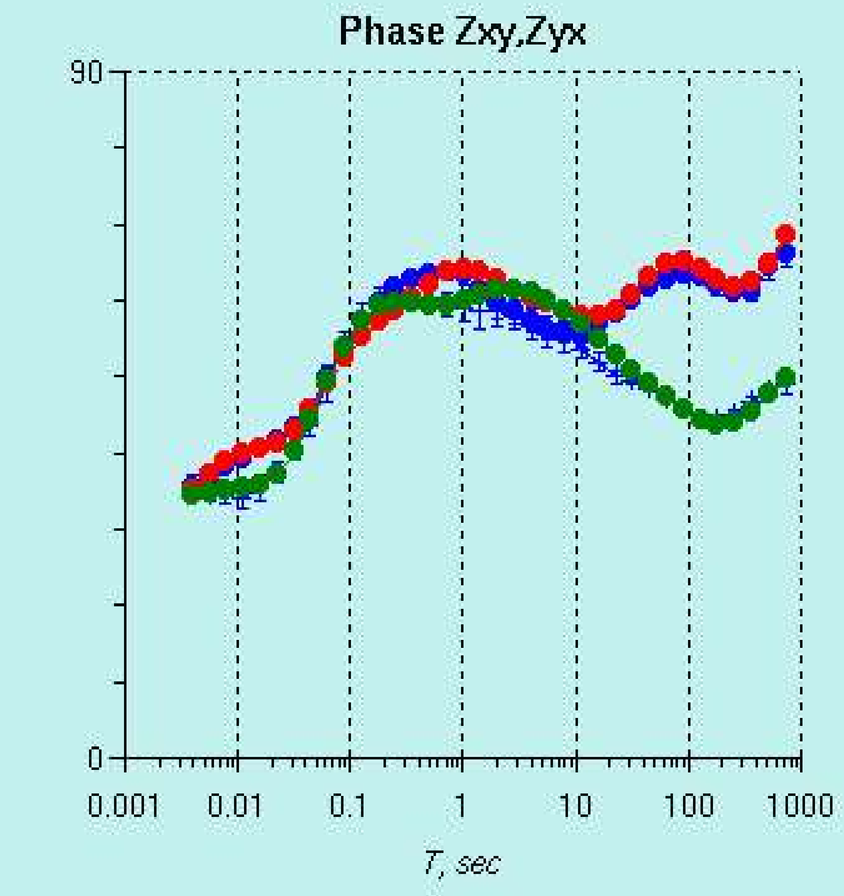


Complete set of RR estimations for AMT site (0.001- 1000 sec). 6 windows (64-262000) and one RR site.



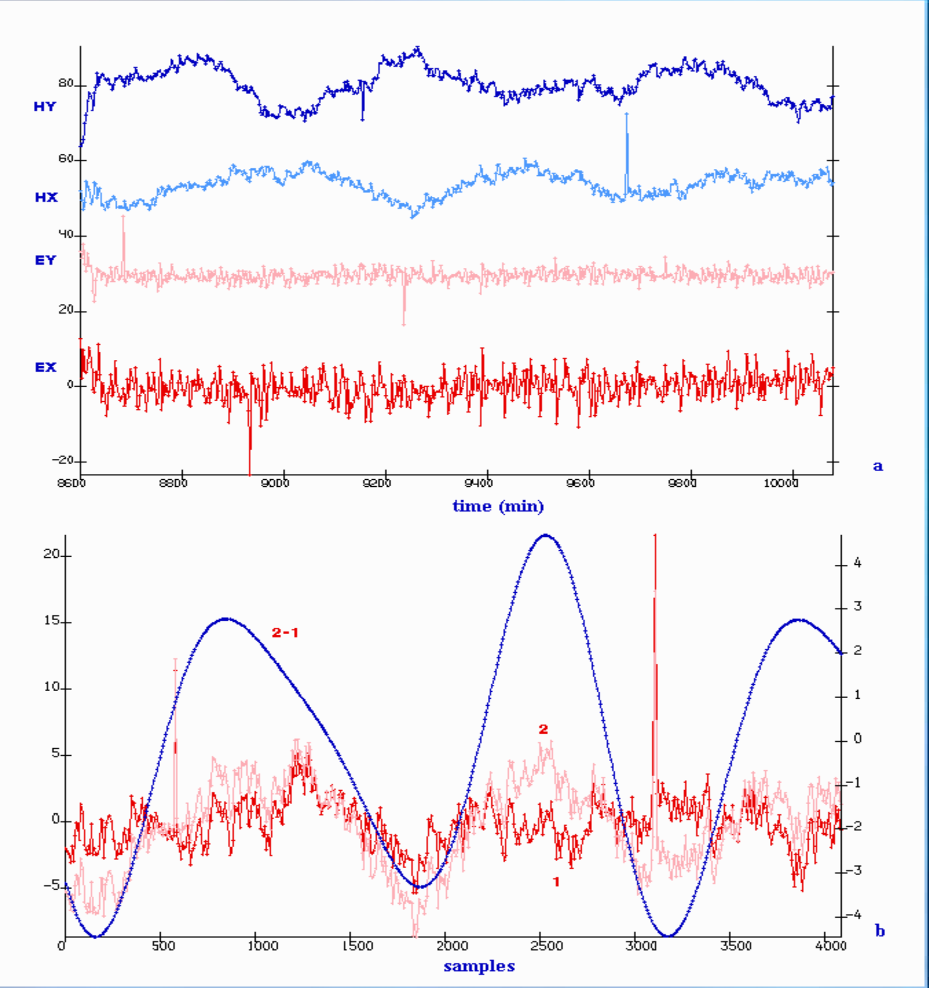
1D consistency check for AMT data shown on previous slides. Blue -final average estimate. Red and green 1D model fit for Zxy and Zyx, respectively.



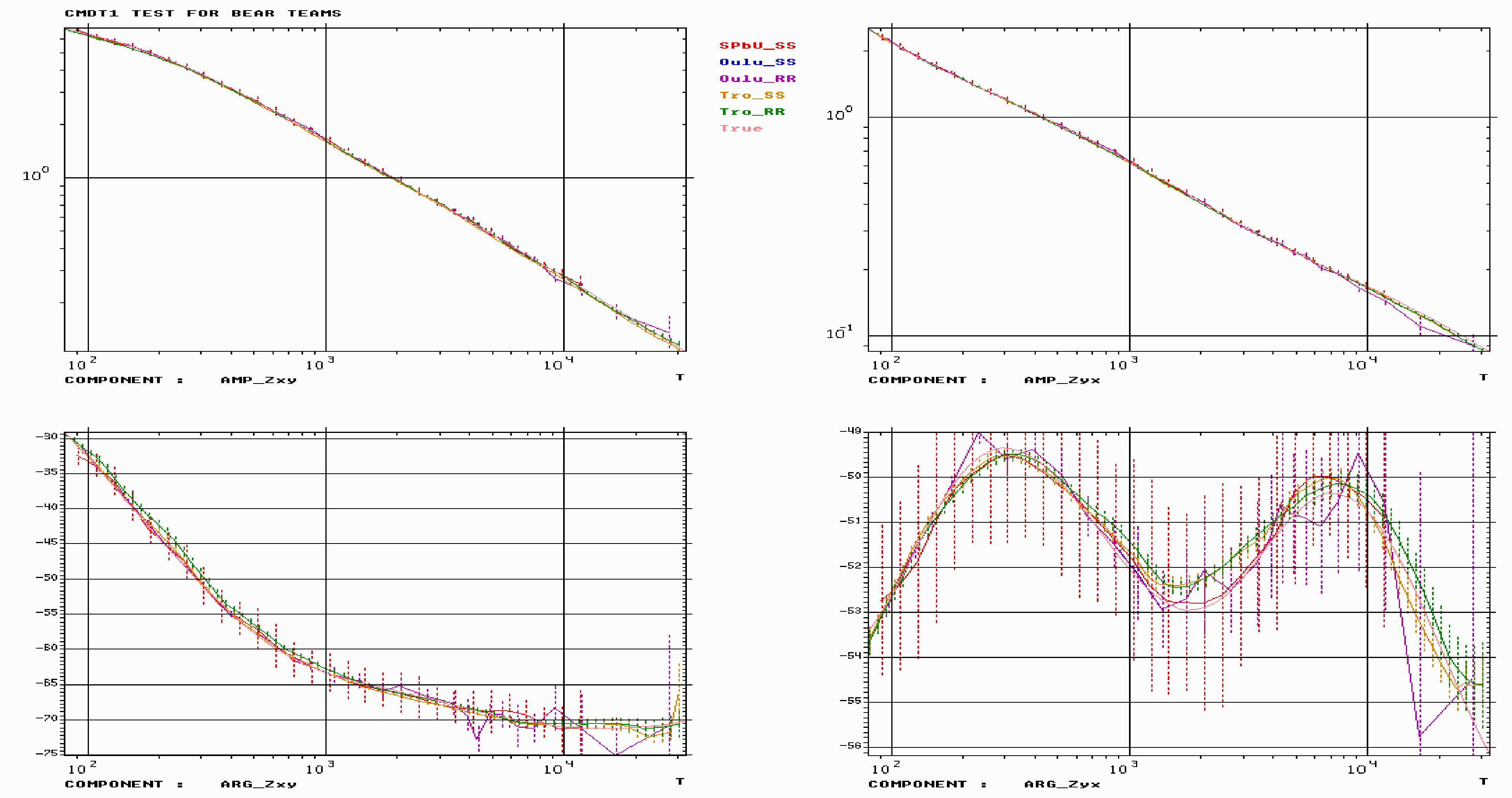


COMDAT synthetic test

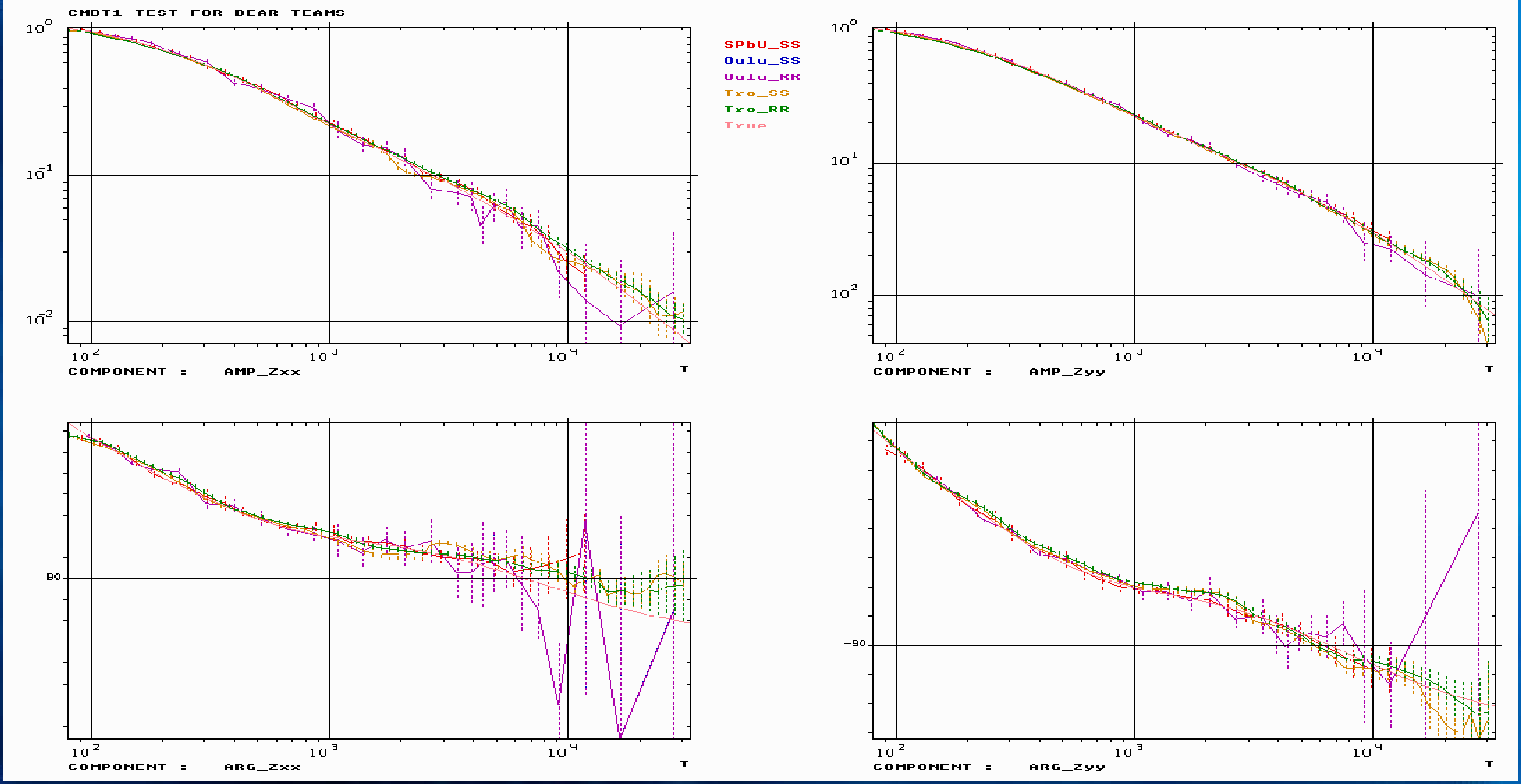
- CMDT3 data set
  - Sample rate 20 sec. 4 components of EM field. The length of time series is 32K samples.
  - Noise in time domain: pulses and strong harmonic noise (first 4000 samples)



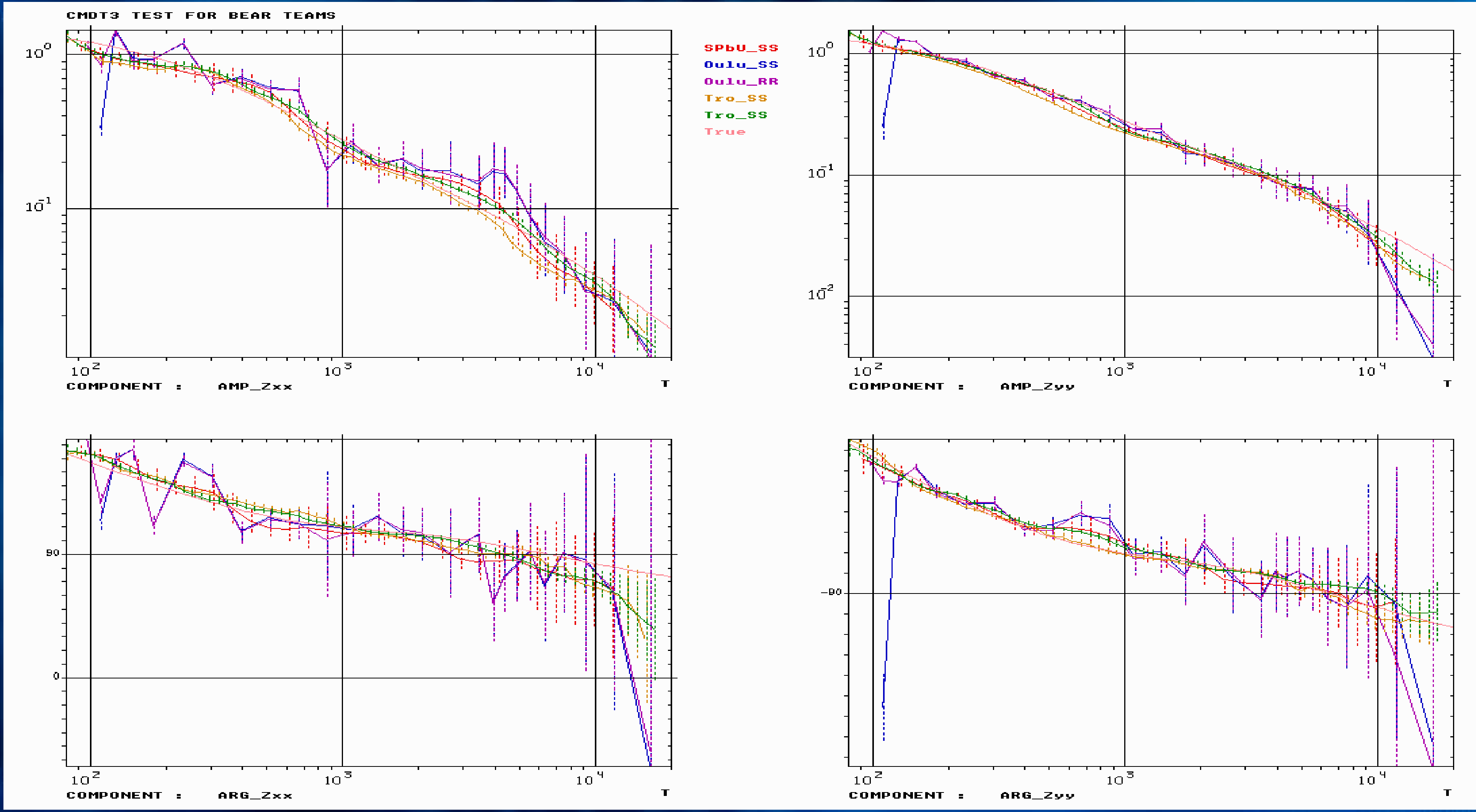
The results of clean data processing. Egbert, Smirnov, Varentsov codes.



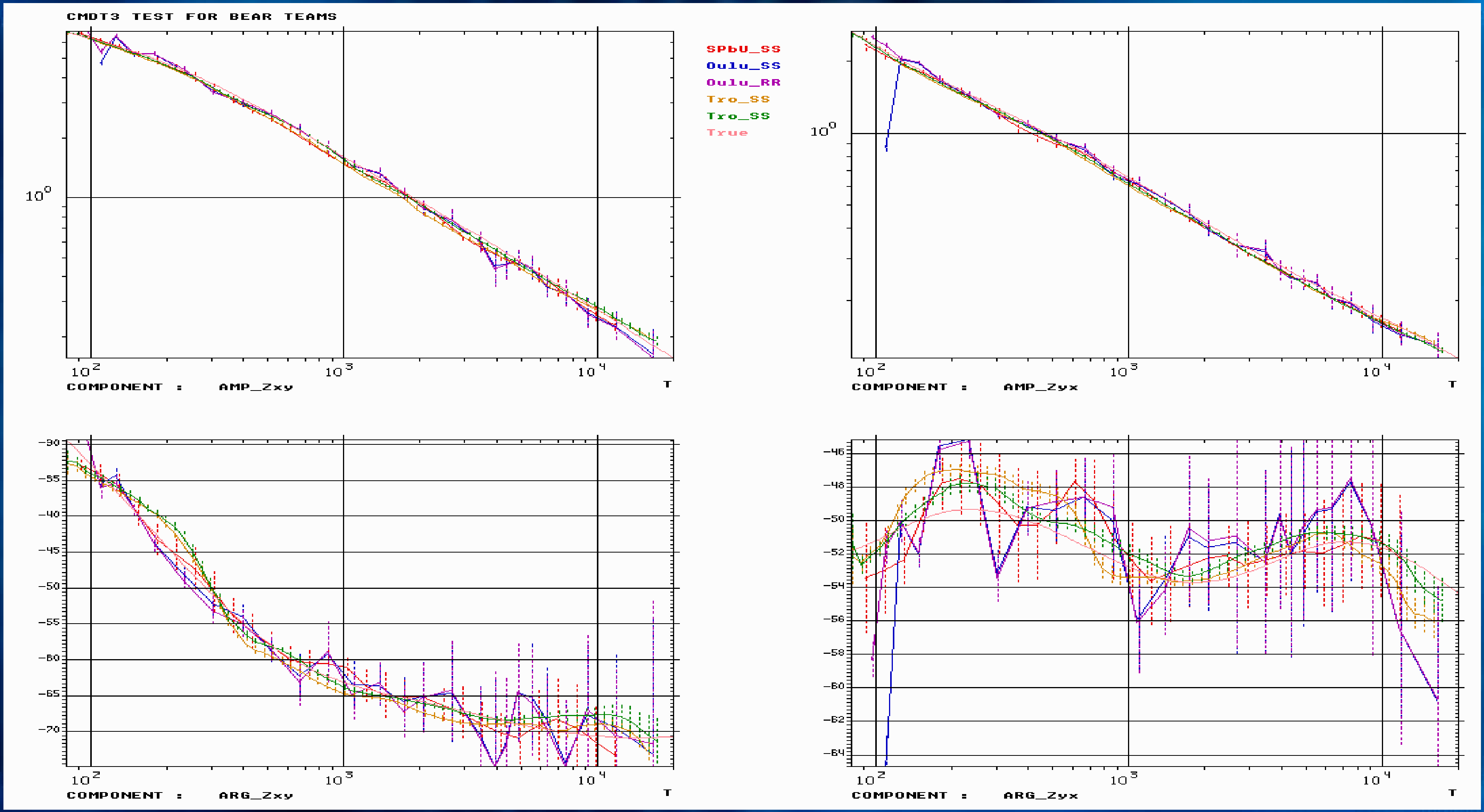
Clean data. Estimates of additional components Zxx, Zyy



Data with added noise. Estimates of additional components Zxx, Zyy



Data with added noise. Estimates of main components Zxy, Zyx



References

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

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

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

Dmitriev, V.I., and M.N. Berdichevsky, 1979. T*he fundamental model of magnetotelluric sounding*, Proc. IEEE, 67, 1033-1044. Durbin, J., 1960. The fitting of time series models, Rev. Int. Stat. Inst., 28, 233-244.

Egbert, G.D., 1997. *Robust multiple station MT data processing*, Geophys. J. Int., 130, 475-496.

Egbert, G.D., and J.R. Booker, 1986. *Robust estimation of geomagnetic transfer functions, Geophys*. J. R. astr. Soc., 87, 173-194.

Egbert, G.D., and D. Livelybrooks, 1996. *Single station magnetotelluric impedance estimation: coherence weighting and the regression M-estimate*, Geophysics, 61, 964-970.

Eisel, M., and G.D. Egbert, 2001. *On the stability of magnetotelluric transfer function estimates and the reliability of their variances*, Geophys. J. Int., 144, 65-82.

Ernst, T., Sokolova, E. Yu., Varentsov, I.M. , Golubev N.G., 2001. *Comparison of two techniques for magnetotelluric data processing using synthetic data sets*, Acta Geophys. Polonica, XLIX, 213 243.

Gamble, T.D., W.M. Goubau, and J. Clarke, 1979. *Error analysis for remote reference magnetotelluric*, Geophysics, 44, 959-968.

Garcia, X., A.D. Chave, and A.G. Jones, 1997. *Robust processing of magnetotelluric data from the auroral zone*, J. Geomagn. Geoelectr., 49, 1451-1468.

Jones, A.G., and J. Spratt, 2002. *A simple method for deriving the uniform field MT responses in auroral zones*, Earth, Planets, Space, 54, 443-450.

Jones, A.G., A.D. Chave, G.D. Egbert, D. Auld, and K. Bahr, 1989. *A comparison of techniques for magnetotelluric response function estimation*, J. Geophys. Res., 94, 14201-14214.

Junge, A., 1996. *Characterization and correction for cultural noise*, Surv. Geophys., 17, 361-391.

Larsen, J.C., R.L. Mackie, A. Manzella, A. Fiordelisi, and S. Rieven, 1996. *Robust smooth magnetotelluric transfer functions*, Geophys. J. Int., 124, 801-819.

Oettinger, G., V. Haak, and J.C. Larsen, 2001. *Noise reduction in magnetotelluric time-series with a new signal-noise separation method and its application to a field experiment in the Saxonian Granulite Massif,* Geophys. J. Int., 146, 659-669.

Sims, W.E., F.X.Bostick, and H.W. Smith, 1971. *The estimation of magnetotelluric impedance tensor elements from measured data*, Geophysics, 36, 938-942.

Smirnov, M. Yu., 2003. *Magnetotelluric data processing with a robust statistical procedure having a high breakdown point*. Geophys. J. Int., 152, 1-7.

Sutarno, D., and K. Vozoff, 1991. *Phase-smoothed robust M-estimation of magnetotelluric impedance functions, Geophysics*, 56, 1999-2007.

Varentsov Iv. M., Sokolova E. Yu.,Martanus E. R. ,Nalivaiko K. V. , and the BEAR Working Group, 2003, *System of Electromagnetic Field Transfer Operators for the BEAR Array of Simultaneous Soundings: Methods and Results,* Izvestiya, Physics of the Solid Earth, Vol. 39, No. 2, pp. 118 148.

Thomson, D.J., and A.D. Chave, 1991. *Jackknifed error estimates for spectra, coherences, and transfer functions*, in Advances in Spectrum Analysis and Array Processing, Vol. 1, S. Haykin (ed.), Englewood Cliffs: Prentice Hall, pp. 58-113.

Textbooks

Bendat, S.J, Piersol, A.G., 2000, *Random Data: Analysis and Measurement Procedures*, New York: John Wiley.

Efron, B., and Tibshirani, R.J., 1993, *An Introduction to the Bootstrap*, New York: Chapman and Hall.

Hampel, F.R., E.M. Ronchetti, P.J. Rousseeuw, and W.A. Stahel, 1986. *Robust Statistics-The approach based on influence functions,* New York: John Wiley.

Huber, P., 1981. *Robust statistics*, New York: John Wiley.

Marple, S.L. Jr., 1987. *Digital Spectral Analysis with applications*, Prentice- Hall, New Jersey.

Rabiner L.R., Gold B., *Theory and Application of Digital Signal Processing*, Englewood Cliffs, NJ: Prentice-Hall, Inc., 1975.