Magnetotelluric Time Series Data Processing of Metronix Instrument Based on Matlab

Fadhil Muddasir[1], M. Lutfi Ismail[2], Yunus Daud[1]

[1]Geothermal Laboratory, Department of Physics, Faculty of Mathematics and Natural Sciences, Universitas Indonesia, Depok Campus, 16425, Indonesia

[2]PT NewQuest Geotechnology, Indonesia

fadhil.muddasir@sci.ui.ac.id

Keywords: MT, Metronix, time series data, robust calculation, apparent resistivity, phase

ABSTRACT

One of the most suitable methods for geothermal exploration is the magnetotellurics (MT) method. This method utilizes electromagnetic waves radiating naturally in the earth's surface to describe the resistivity structure of subsurface with a wide frequency range. There are various instruments in MT measurement such as Phoenix Instrument and Metronix Instrument. ADU-07e (Analog Digital Unit) is an instrument of Magnetotelluric method which is from the Metronix. To perform the data processing of this instrument, it is needed software from Metronix itself such as Mapros. We will use the Matlab programming algorithm for data processing of Metronix MT data to make it more friendly and add feature for handling noise such as robust calculation. In this study, it is used time series format (.ats) of the Metronix instrument. The reduction from outlier effect on impedance data is handled by robust calculation. The results of this study indicated that curve resistivity from data processing using Matlab program with robust calculation feature showed the curve smoother compared with the Mapros results. Two-dimensional inversion model showed a significant result in the Base of Conductor determination for geothermal system.

INTRODUCTION

The Magnetotelluric method utilizes electromagnetic waves that radiating naturally in the earth's surface to describe the subsurface resistivity structure with a wide frequency range. Magnetotelluric method measures the electrical field and the Earth's magnetic field that occurred naturally in the earth's surface (Vozoff, 1972). The source can originate from lightning activity for more than 1 Hz frequency and the interaction between the solar wind and the Earth's magnetosphere layer with a lower frequency of 1 Hz. The wide range in frequencies used means that this method can overcome the overburden layers of conductive and has deep penetration. This method measures the electrical field and magnetic field in two directions which is perpendicular to each other, so it can provide important information related to electrical anisotrophy in certain areas (Daud, 2010). Because of this characteristic of the MT method, it makes more effective and powerfully to be used in geothermal exploration.

In MT method, the horizontal component of the electrical field and the three components of the magnetic field are measured at the surface. The impedance can be obtained by measuring the variation of the electrical field and magnetic field simultaneously.

There are various instruments to measure MT such as Phoenix Instrument and Metronix Instrument. Currently, the industry of geothermal exploration is common widely used Phoenix instruments and software associated with the instrument. ADU-07e (Analog Digital Unit) is an instrument of Magnetotelluric method which comes from Metronix Geophysics. To perform the data processing of this instrument, it needs software from Metronix like Mapros. One of the disadvantages of the software is not user friendly and practical to be used, especially in handling noise. During this time, Metronix MT data processing software has not been developed. Therefore, The author will use the Matlab programming algorithm for data processing of the data MT's Metronix and adds feature for handling noise such as robust calculation. This research was carried out with an idea based on previous research conducted by Heditama about MT data processing which is using time series format (.ts) from Phoenix instrument to obtain apparent resistivity and phase (Heditama, 2011), While it is using time series format (.ats) of the instrument Metronix in this reseacrch. The results of data processing MT in this research is the plot curve spectra, apparent resistivity, phase, and coherence. The impedance data from the processing will be made a visualization using the 2D inversion software so it will be compared with the modelling results of processing using Mapros software.

TIME SERIES PROCESSING

Magnetotelluric method is a geophysical exploration techniques that involve the relationship between horizontal components of the magnetic field and electric field to determine variations in electrical conductivity below the earth's surface (Cagniard, 1953). Electromagnetic fields (EM) are caused by various physical processes are quite complex so that the frequency spectrum is very wide (10^{-5} Hz to 10^4 Hz). At sufficiently low frequencies (less than 1 Hz), solar wind containing electrically charged particles interact with permanent Earth's magnetic field, causing EM field variations. Magnetotelluric data processing involves the concept of electromagnetic waves, time series analysis and linear system theory impedance interpreted with numerical modeling scheme. The reduction in data processing MT is done to eliminate noise participating measured at the time of data acquisition.

Time Series data were obtained from the acquisition in the field is a trailer signal magnetic field in the x, y, and z axis, and the electric field in the x and y axis for each sounding to see the characteristics of the signal and noise were recorded. Time series data is transformed into the frequency domain spectrum through Fourier transformation. In addition, through the time series data is done Fourier coefficient calculation. After that the data is included calibration parame-

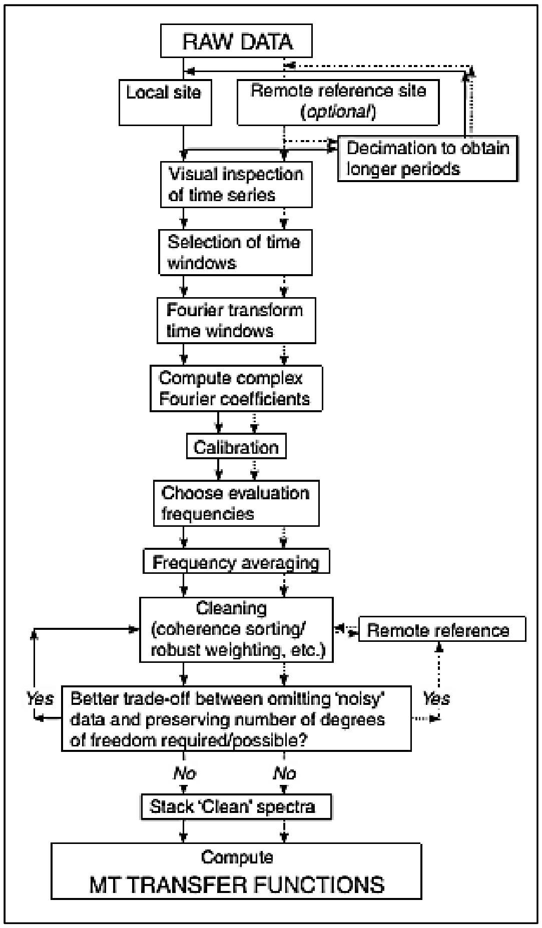


Figure 1. Flowchart The Steps Involved in MT Data Processing (Simpson and Bahr, 2005)

ters to avoid sampling errors caused by limitations in measurement.

Once the frequency domain data is combined with the calibration data, the next step is to determine the matrix spectra and averaging frequency. At this stage the data can be analyzed using coherency to estimate the quality. If the value of coherence is still undervalued, then the data needs to be stacking to increase the signal to noise ratio value, then it is formed in the transfer function to be applied in a variety of interpretative data such as apparent resistivity and phase curves, tipper, induction vector, etc. In Figure 1 in more detail the stages of data processing MT is described.

MAGNETOTELLURIC ON GEOTHERMAL SYSTEM

In geothermal system is a process flow of heat from the heat source to the heat sink normally consists of several components, including cap rocks, reservoir, heat source, and the fluid in the reservoir. Cap rocks is generally a clay layer that is low permeable reservoir located above. This layer serves as closing the fluid flow and heat the reservoir zone that flows as a result of convection processes. Still, this layer can not prevent the flow of heat to the earth's surface that occur by conduction. However, despite the amount of heat conduction is very small compared with the flow of heat escape through the process of convection. Cap rock layer is a rock alteration of the result of hydrothermal alteration processes that occur in a geothermal system called clay cap. Hydrothermal alteration process can produce several types of clay cap hanging with

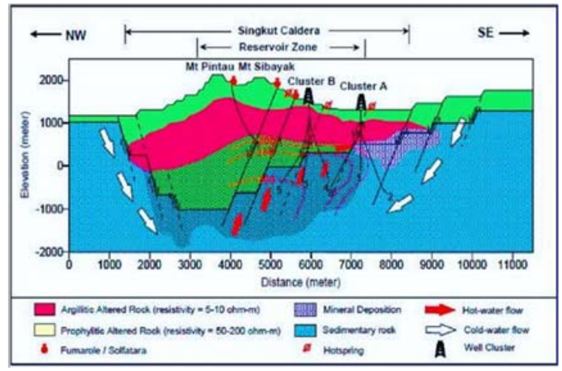


Figure 2. Geothermal System (Daud, 2002)

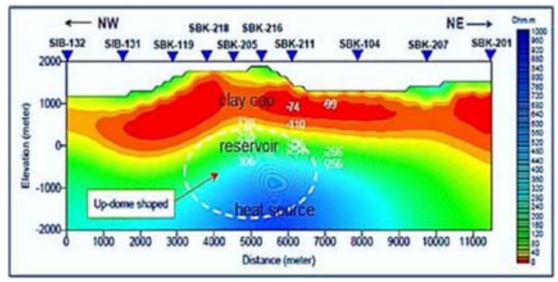


Figure 3. MT Response on Geothermal System (Daud, 2002)

clay rock temperature at the time of the process. Layer smectite clay type and mix of types of smectite-illite clay formed in the temperature range < 200°C. Reservoir which became the accumulation of heat energy into the target zone for storing the fluid in the form of hot water or high-pressure steam. However, in the reservoir rock porosity and permeability ideally have a great natural so it can accommodate a lot of fluids and fluid flow in it can work well.

Heat source in the system is hot rock usually derived from the intrusion of magma. This magma intrusion can creep closer to the surface through fractures which appeared on crystalline bedrock (basement). Figure 2 shows an example of a geothermal system.

MT response on geothermal systems generally show a distribution layer with resistivity 1 – 15 Ωm as clay cap, a layer with resistivity 30 – 50 Ωm under the clay cap as a reservoir, and a layer of above 80 Ωm as a heat source. Figure 3 shows the response of MT on the geothermal system. To reinforce the indications of geothermal system below the earth's surface, generally the response MT will form an up-dome shaped arising from the flow and thermal gradients of heat source that propagate radially.

MATLAB PROGRAM IN MT DATA PROCESSING

In this research, the authors do steps in the developing of MT data processing program from Metronix instrument with additional features for robust calculation using Matlab programming language with the flowchart is indicated by Figure 4. Beside of featuring robust calculation, the program added feature analysis functions such as coherence and autocorrelation. ATS data conversion into ASCII data was done using binary data decoding and then the data is encoded in the form of ASCII Hexadecimal with format *.dat file. The conversion process is done to facilitate the process of further analysis and data processing.

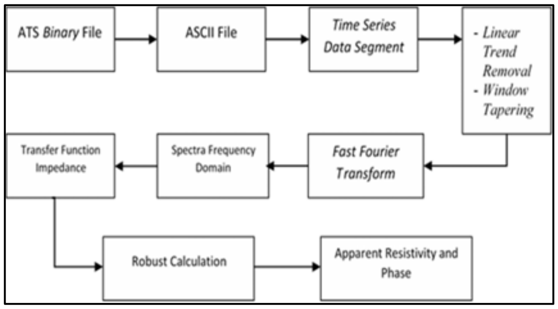


Figure 4. Program Flowchart

The next process is the linear trend removal which aims to eliminate the linear trend in time series data which should be the trigonometric functions. The effect of this linear trend is able to bring spectral leakage in the frequency domain data obtained from the Fourier transform. The process which is performed on linear trend removal is to change the average gradient value of time series data to be zero. This process uses an external function in Matlab, *y = detrend (x)*.

Then, processing continues with segmentation of data where the data measured in the time of measurement data is divided into segments with a maximum number is approximately the length of data divided by 1024 or 81920 for long span data. This was done to simplify the elimination of noise and Fourier transformation, especially in the selection of time series data. The disadvantages of this segmentation process is able to lower the resolution and accuracy of the frequency domain data when the time series was transformed. To overcome this, the data will be convoluted with a tapering window function before Fourier transformed. Window functions used in program Matlab is Hanning, since this function has advantages in terms of increasing the resolution and accuracy of data from the Fourier transformation compared to other window functions such as Rectangular and Nuttal.

Transforming data from the time domain to the frequency domain is used Fourier transformation, particularly, Fast Fourier Transform (FFT) so that the process of transformation is done fast and efficiently. Fast Fourier Transform (FFT) is a computational method that is capable of handling discrete Fourier transform of the data with large numbers efficiently. Efficiency is mainly from its ability to take advantage of the properties contained in the periodic functions of sine and cosine.

Function used in the Matlab algorithm is *Y = fft (x)* and *y = ifft (X)* implement the transform and inverse transform pair given by the vector of length N (The Mathworks, 2014) are shown in the following equation

$$X(k) = \sum_{j=1}^{N} x(j)\,\omega_N^{(j-1)(k-1)} \qquad (1)$$

$$x(j) = \left(\tfrac{1}{N}\right)\sum_{k=1}^{N} X(k)\,\omega_N^{-(j-1)(k-1)} \qquad (2)$$

where $\omega_N = e^{(-2\pi i)/N}$

The results of this calculation is then processed into the impedance transfer function in Matlab using the following equation

$$Z_{ij} = \frac{\langle \tilde{H}_i E_i^* \rangle \langle \tilde{E}_i \tilde{H}_j^* \rangle - \langle \tilde{H}_i \tilde{H}_j^* \rangle \langle \tilde{E}_i \tilde{E}_i^* \rangle}{\langle \tilde{H}_i \tilde{E}_i^* \rangle \langle \tilde{H}_j \tilde{H}_j^* \rangle - \langle \tilde{H}_i \tilde{H}_j^* \rangle \langle \tilde{H}_j \tilde{E}_i^* \rangle} \qquad (3)$$

then the impedance value from segments of data is estimate by robust calculation. The results of impedance from robust calculations will then be calculated apparent resistivity and phase. The process of calculating the apparent resistivity and phase in Matlab using the following equation

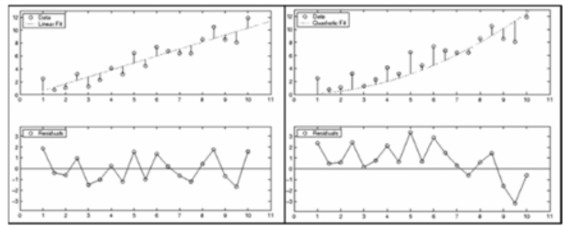


Figure 5. Residual Analysis on Linear Data (Left) and Quadratic (Right) (The Mathworks, 2014)

$$\rho_{aij} = \frac{\mu_0}{\omega}\left|Z_{ij}\right|^2 \qquad (4)$$

with its error,

$$\Delta\rho_{aij} = 2\rho_{aij}\frac{\Delta Z_{ij}}{\left|Z_{ij}\right|} \qquad (5)$$

and impedance phase,

$$\varphi_{ij} = arctan\left(\frac{Im(Z_{ij})}{Re(Z_{ij})}\right) \qquad (6)$$

with its error,

$$\Delta\varphi_{ij} = \frac{\Delta Z_{ij}}{\left|Z_{ij}\right|} \qquad (7)$$

it can be necessary to multiply Z with 1000 before calculating the absolute value for the case of the electric field has the unit mV / km (Friedrichs, 2007).

ROBUST CALCULATION

The focus in this research is the addition of robust calculations on the data processing program of MT Metronix instruments. In this case, the author use the Matlab programming algorithms derived from Garcia's research on smoothing the data using the robust data in one and higher dimension with the missing values on data (Garcia, 2009). In this case, the author make modifications to the algorithm to be applied in estimating the value of the MT impedance contaminated by outlier in the data to get the curve apparent resistivity and phase more smooth than without using a robust calculation.

In regression analysis, it is assumed that the residuals of the data follow a normal distribution / gaussian. The residual value is the difference between the response of the data with the fit result of data in the form of predicted value or model of the data. The residuals appear randomly scattered around zero indicating that the model describes the data well (The Mathworks, 2014). Figure 5 above shows the process of analyzing the residual of data linear and quadratic.

Firstly it is performed outlier identification which are contained in the data, the residual value of the outlier data points will have a high value. In the estimation of data using least square, residual value of the outlier that contaminate the data has a value of high leverage. Leverage value in the statistics showed a measure of influence (between 0 to 1) from the data points on a fitting model due to its location in the space of the inputs. More precisely, the points which are far removed from the main body of points will have high leverage (Chatterjee and Hadi, 1986).

That is causing the least squares estimation will cause outlier that contaminate the data has a high effect on fitting the model. To minimize the effect of outliers, can be done by downweight the influence of outlier data points, as often used in robust regression.

Several functions that are used in weighting process is performed on robust regression. Most commonly used functions are bisquare weight function. This function is expressed by

$$W_i = \begin{cases} \left(1 - \left(\frac{r_i}{6MAD}\right)^2\right)^2 & ,if\ |r_i| < 6MAD \\ 0 & ,if\ |r_i| \geq 6MAD, \end{cases} \quad (8)$$

with

$$r_i = y_i - \hat{y}_i \quad (9)$$

where r_i is residual from *i-th* observation and $\hat{\sigma}$ is robust estimator for standard deviation of residual given by 1.4826 MAD, where MAD denotes the Median Absolute Deviation (Rousseuw *and* Croux, 1993), given by

$$MAD = median\ |x_i - median\ (x)| \quad (10)$$

The use of bisquare weights combined with residual analysis provides robust data more smooth. The value of the data that has been downweight will be performed an interpolation on the data to get robust fitting results more smooth in estimating MT impedance data.

TEST ON SYNTHETIC DATA

In this research, the performance tests conducted by the author on synthetic data that is contaminated with random noise added to the data by using a synthetic signal with wave equation given by $y = A \sin \omega t$ on frequency of 2 Hz. So it will decrease the value of the autocorrelation that is used as parameters to test the synthetic signal.

On Figure 6, it shows that the signal for 7 seconds has random noise that contaminates 10% of data. Then these signals has done segmentation to obtain n sample segments of data. In this synthetic signal, the author divides the signal into 7 segments as shown in Figure 7. After the segmentation of the signal, the signal is transformed to the frequency domain using Fourier transformation. As shown in Figure 8. It shows the first segment of the signal is contaminated by noise so that when it is transformed into the frequency domain it look the spectral leakage value on the signal indicated with a blue circle in the first segment compared with the average spectrum indicated by red dots on the other segments. After all the signals are transformed into the frequency domain, robust calculations performed for the data in each frequency on all segments of the data. Robust calculations are used to minimize the effect of outliers in the data due to the noise added to the signal.

Then after robust calculation, calculation of the median value from the distribution of data for each segment is applied and standard deviation from the median of the data. Then, it is set up the frequency domain data of robust calculation result and doing inversion of Fourier transform to analyze the effect of robust calculation on the results when the signal is reversed to the time domain.

Figure 9 shows frequency domain data after robust calculation. This data will be reversed to time domain to obtain signal after robust calculation by minimalize the effect of outlier. Figure 10 shows the signal after robust calculation. For the purposes of analysis, the authors used the autocorrelation parameter that is shown in Figure 11 to analyze the autocorrelation signal before and after the robust calculation.

Table 1. Comparation Result for Synthetic Signal Test

Noise Contaminated	Robust Calculation	Autocorrelation
10% noise	Good	Increased
20% noise	Good	Increased
40% noise	Good	Increased
50% noise	Bad	Increased Slightly
100% noise	Bad	Not Increased
Harmonic Power Line	Not Filtered	Not Increased

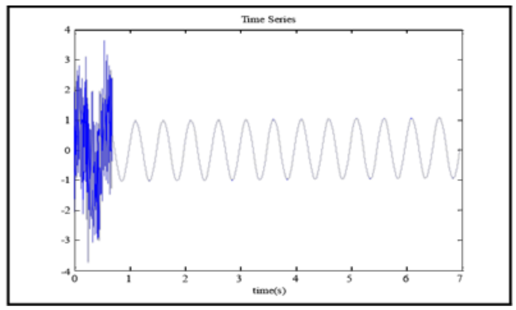


Figure 6. Synthetic Signal Contamined by 10% Noise

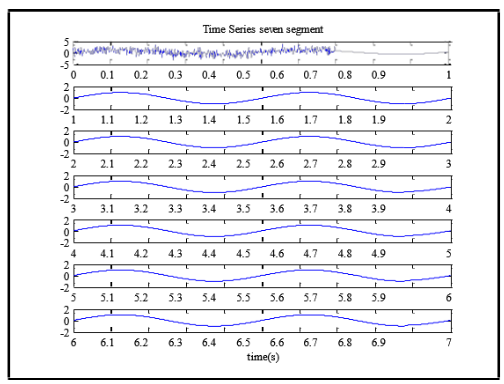


Figure 7. Synthetic Signal Segments

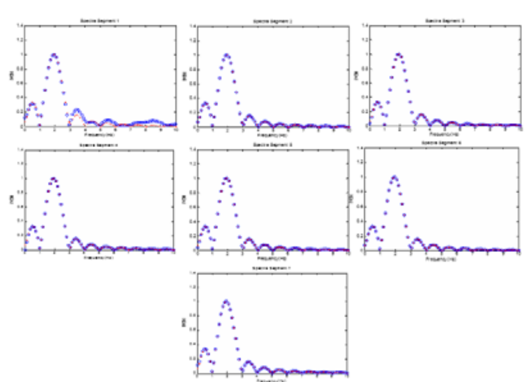


Figure 8. Spectra Frequency Domain Segments

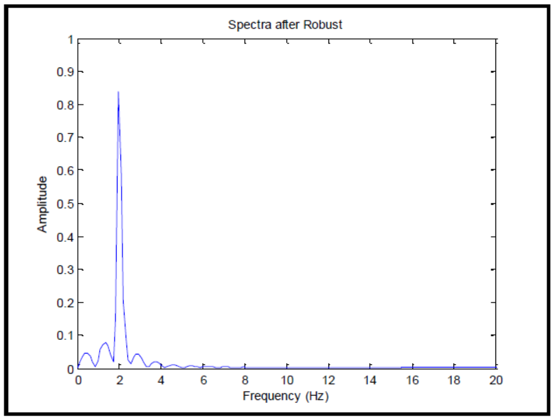


Figure 9. Spectra after Robust Calculation

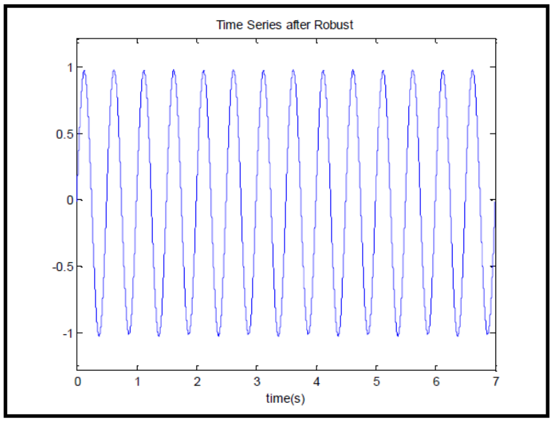


Figure 10. Synthetic Signal after Robust Calculation

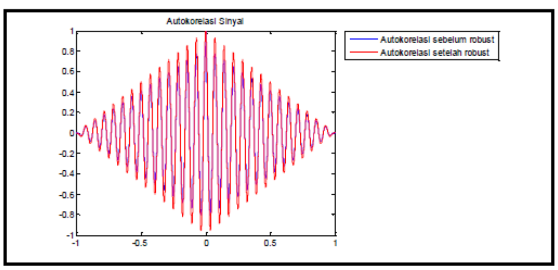
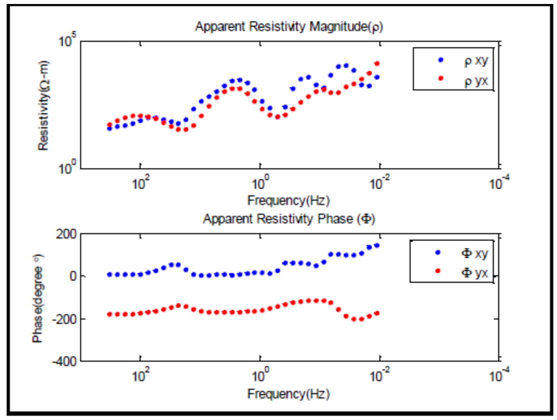
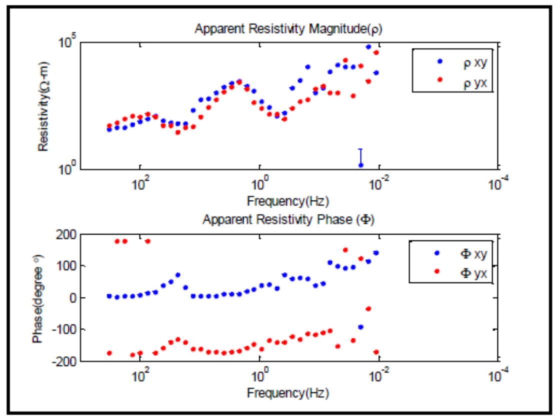


Figure 11. Autocorrelation of Signal Before (Blue) and After (Red) Robust Calculation

From the results, it can be seen that the synthetic signal contaminated 10% noise before robust calculation, the autocorrelation value is lower than the signal that has been carried out robust calculation. After that, the synthetic signal tested on noise contaminated data by 20%, 40%, 50%, 100%, and Harmonic Power Line noise. From the results, several tests perform in the contaminated synthetic signal by noise and do the visual inspection and also analysis of the autocorrelation parameters, are shown in Table 1. From Table 1, it can be concluded that the robust calculations process are applied to a synthetic signal will yield significant results in contaminated signal noise below 50%. It can be seen from the calculation robust results that can filter out noise from the data. As for the synthetic contaminated signal noise by 50% and more and also contaminated by Harmonic Power Line noise will not have a significant effect due to robust calculations only filter out noise that is considered as an outlier of the data distribution. Noise contaminating the signal at 50% and more of the signal length or with low signal to noise ratio cannot be filtered by robust calculation.



(a)



(b)

Figure 12. Apparent Resistivity and Phase Station TR0043, **(a)** Matlab with robust calculation and **(b)** Mapros without robust calculation

While Harmonic Power Line noise is coherent and harmonic to signal so robust calculations cannot filter the noise because it cannot be identified as an outlier.

REAL MT DATA APPLICATION

The authors also tested the program on real data MT. Robust calculation process on the real data is performed on the frequency domain impedance data. Data MT consisted of nine stations in one line measurement. The stations are TR0043, TR0045, TR0046, TR0047, TR0048, TR0049, TR0050, TR0051, and TR0052.

In Figure 12 is shown the apparent resistivity and phase that were performed using robust calculations and without robust at the station TR0043. The result of the apparent resistivity curve data using robust calculations showed the apparent resistivity curve more smooth, especially at frequencies over 1 Hz. On the data at frequencies around 0.03 Hz, there are outliers shown far from the trend curve on the data without robust calculation. While the data with robust calculation, the influence of outliers on these trends can be minimized and obtained more smooth curves. In the data below a frequency of 1 Hz obtained curves slightly different, especially in the deadband frequency. But on overall, curve of apparent resistivity and phase using robust calculations smoother than without robust calculation.

In the comparison result from the other stations of MT, for data MT that occurred at the deadband frequency still not produced the curve smoother and trends on the curve resistivity apparent at frequencies below 1 Hz occur a considerable margin between the data processing results with robust calculation compared with data processing without robust calculation. This is likely due to the more of noise than signal in the data and electromagnetic wave attenuation below the earth's surface at the deadband frequency resulting signal attenuation and the resulting data is less accurate.

INVERSION AND VISUALIZATION

Two-dimensional inversion conducted using Non-Linear Conjugate Gradient (NLCG) algoritm to see the effect of robust calculation to the inversion models. Depth modeling and inversion limited to ~3000 meter. The results of Invariant mode inversion and modeling of the data processed by Matlab shown at Figure 13, while the data is processed by Mapros shown in Figure 14.

Then the comparison of the model inversion result with Matlab and Mapros processing above shows that the inversion result using the input of data that is processed using Matlab with RMS error 24.1% and Mapros with 25.3%. The model shows the up-dome shaped form which is the Base of Conductor of the geothermal system. Updomed- shape is a geological structure arising from the heat source the earth's subsurface and causing a radial gradient of temperature, which in the zone around the heat source. This shows that robust calculation program that has made a significant contribution in the inversion and MT modeling.

CONCLUSIONS

MT data processing procedures using program based on Matlab generally include data segmentation step, digital filter process and window tapering to increase the signal resolution and then Fourier transform from the time domain data to the frequency domain. Impedance is calculated using the Transfer Function and featured by robust calculation to estimate it. Furthermore obtained apparent resistivity and phase from impedance data.

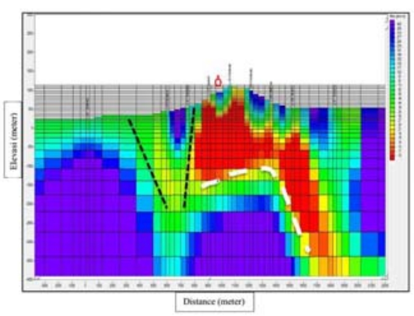


Figure 13. 2D Inversion Result from Data that is Processed by Matlab

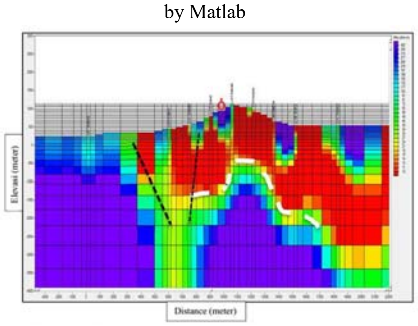


Figure 14. 2D Inversion Result from Data that is Processed by Mapros

The results of the comparison of data processing using Matlab featured by robust calculation and data processing using Mapros software without robust calculation shows that the effect could generate apparent resistivity and phase curve smoother in data processing using Matlab after minimizing outlier effect on impedance data compared with the results of data processing using Mapros software.

On further research needs to be studied again the calibration process of the data on the Metronix instrument and the software Mapros to be applied to the development of Matlab-based program and adding Digital Filter to filter out coherent noise affected in signal, remote reference, and the cross power selection in the frequency domain after robust calculations to improve data quality, especially in the deadband frequency.

REFERENCES

Cagniard, L. (1953). Basic Theory of The Magnetotelluric Method of Geophysical Prospecting, Geophysics, 18, 605-635

Chatterjee, S., Hadi, A.S. (1986). Influential Observations, High Leverage Points, and Outliers in Linear Regression. Statistical Science 1, 379-393.

Daud, Y. (2010). Electromagnetic Method: Success Story in Geothermal Exploration & Possibility for Hydrocarbon Exploration. College Textbook, Depok.

Daud, Y., Sudarman, S., & Ushijima, K. (2002). Imaging Reservoir Structure of the Sibayak Geothermal Field (Indonesia) Based on Magnetotelluric Observations. Geothermal Resources Council Transactions, Vol.26, September 22-25.

Friedrichs, B. (2007). MAPROS, ver 0.87b freeware. Metronix GmbH. Braunschweig, Germany.

Garcia, D. (2009). Robust Smoothing of Gridded Data in One and Higher Dimensions with Missing Values. Computational Statistics and data Analysis, 1167-1178.

Heditama, D.M. (2011). Pemrosesan Data Time Series Pada Metode Magnetotelurik (MT) Menjadi Data Resistivitas Semu Dan Fase Menggunakan MATLAB. Skripsi S-1 Jurusan Fisika, Universitas Indonesia.

Rousseeuw, P.J., Croux, C. (1993). Alternatives to The Median Absolute Deviation. Journal of The American Statistical Association 88, 1273-1283.

Simpson, F., and Bahr, K. (2005). Practical Magnetotellurics. Cambridge: Cambridge University Press. The Math Works, Inc. (2014). Retrieved June 14, 2016. From The Math Works Web site: http://www.mathworks.com/help

Vozoff, K. (1972). The Magnetotelluric Method in The Exploration of Sedimentary Basins, Geophysics, 37, 98 – 141.