Estimating Noise Levels in AEM Data

Andy Green
OTBC Pty. Ltd, Australia
greenaa@ozemail.com.au

Richard Lane
Geoscience Australia, Australia
richard.lane@ga.gov.au

SUMMARY

This paper reports the results of analysing AEM data where the same flight line has been flown repeatedly to monitor system performance.

When these data are corrected for the effects of variable survey altitude they provide a good measure of the reproducibility of AEM data. The analysis has been conducted for both frequency and time domain systems and shows that, where the area is even moderately conductive, multiplicative errors will provide the dominant source of noise for AEM surveys. Errors typically have standard deviations of 2 % and can easily induce fluctuations of 10%. Moreover, because of their geometric origins, these errors are highly correlated between channels. Frequency domain HEM and Z component time domain systems produce errors that are largely unrelated to time/frequency/conductivity. This is not the case for X component data from asymmetric systems.

Key words: Noise, AEM

INTRODUCTION

An accurate understanding of the noise characteristics of an AEM system is necessary to perform unbiased manual interpretation and to correctly weight the data in most detection and inversion algorithms. Unfortunately, there is very little information on the realistic, in-flight noise levels of the available AEM systems. The values commonly published are often for special situations such as ground-based measurements of electronic noise or measurements from straight and level flight at high altitude.

This conservative attitude to defining noise levels is unsurprising because, when more complicated sources of variability are considered, it is hard to fairly distinguish between noise and other systematic sources of variability that could, in principle, be monitored and incorporated into the data processing strategy. At the moment the only source of variability that is routinely treated in this way is the survey altitude. However there are many other sources of geometric variability that are currently unmonitored and uncompensated. These often provide the major source of along-line variability and, at least until they are fully compensated, should be treated as noise. In addition to this geometry-related noise, other noise sources, such as drift in fixed geometry systems or primary field removal errors in variable geometry systems, are difficult to quantify.

In this paper we report the results of analysing data where the same flight line has been flown repeatedly. Two surveys are discussed here. One survey with a 6 frequency Dighem Resolve system had 27 repeat lines, each about 5 km long. In the Tempest time domain data set, we have 11 repeated lines around 11 km in length, with 15 channels of both X and Z components.

A MODEL FOR NOISE

Let us assume we have repeat flight lines where each along-line sample is in the same location for all lines. Then, for any given channel, we have data $X_{l,i}$ where l is line number and i is sample number. In the case of time domain data these X's are real quantities, while for frequency domain data they are complex.

The most general linear model for noise in these data would allow for both additive and multiplicative errors. Thus,

$$X_{l,i} = E_{l,i}R_i + A_{l,i}$$

Here the E's are multiplicative errors and A's are additive errors that are functions of line and sample. The R's are the true ground responses uncorrupted by error. They are not a function of line number.

Analysis of this model is complicated and, initially, we work with simplified versions where the error is either additive (E = 1) or multiplicative (A = 0). Taking the additive case first our error model will be

$$X_{l,i} = R_i + A_{l,i}$$

If we then compute the three averages over lines (l), locations (i) and both together:

$$X_{\bullet,i} = A_{\bullet,i} + R_i$$
$$X_{l,\bullet} = A_{l,\bullet} + R_{\bullet}$$
$$X_{\bullet,\bullet} = A_{\bullet,\bullet} + R_{\bullet}$$

we can calculate a residual

$$D_{l,i} = X_{l,i} - X_{\bullet,i} - X_{l,\bullet} + X_{\bullet,\bullet}$$
$$= A_{l,i} - A_{\bullet,i} - A_{l,\bullet} + A_{\bullet,\bullet}$$

that is independent of ground response terms.

It should be noted that the $A_{l,\bullet}$ term will contain any "drift" errors that are constant over the whole flight line. This means that our error estimates will not include an estimate for this type of noise.

If the errors are multiplicative we can achieve an equivalent result by taking logs first and then applying the same procedure. Thus, if we let lower case letters represent the logarithm of a quantity (e.g. $\log(D_{i,l}) = d_{i,l}$) we have

$$d_{l,i} = \varepsilon_{l,i} - \varepsilon_{\bullet,i} - \varepsilon_{l,\bullet} + \varepsilon_{\bullet,\bullet}$$

If we have enough samples we might expect that the averages of these noise terms to be either zero or one, and thus the residuals become estimates for the noise. That is for additive and multiplicative errors respectively we have,

$$D_{l,i} \approx A_{l,i} \text{ and } d_{l,i} \approx \varepsilon_{l,i}$$

In the case of the complex, frequency domain data additive errors should be processed on the In-phase and Quadrature and multiplicative errors on the log(amplitude) and phase.

If we can assume that all the error terms are uncorrelated with each other and with the ground response ground response (R_i) then we can make some more general conclusions with regard to the combined multiplicative and additive model. As we shall see these assumptions are probably only valid for some types of system geometry but under this model, if we compute the residuals as an additive model we can obtain:

$$D_{l,i} = (E_{l,i} - 1)R_i + A_{l,i}$$

and thus expressions for the variance of the errors,

$$Var(D_{l,i}) = Var(E_{l,i} - 1)\frac{1}{N_i}\sum_i R_i^2 + Var(A_{l,i})$$

This gives us a way of estimating both multiplicative and additive noise under these assumptions.

PROCESSING

The processing discussed here uses the 6 Dighem Resolve frequencies as an example to illustrate the analysis of multiplicative errors.

Preliminary processing is necessary to register the repeat lines to a common flight line. This common line was estimated by a least squares fitting procedure that calculates a "2-way" least squares fit on the x,y data by minimizing the normal deviates from the line, York (1966).

If we now look at the normal deviates of each of the repeat lines from this common line, we get the plot shown in Figure 1 that illustrates the accuracy with which the 27 lines were flown.

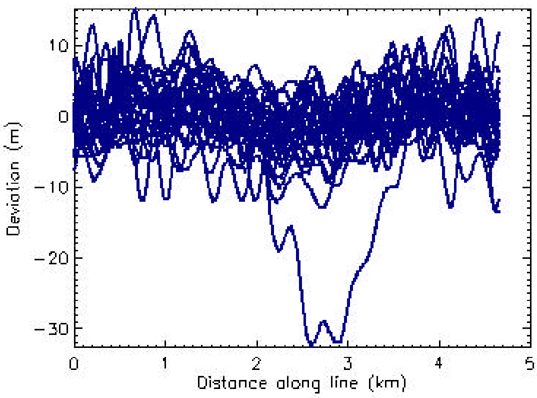

Figure 1. Deviation from the common line for 27 Dighem flight lines

Once a common line is established we can resample the variables in which we are interested onto a common set of samples. Of course this new, resampled data covers only those sections of the data that are common to all lines. All subsequent discussion will focus on these resampled data rather than the raw data.

The noise values require some form of qualification regarding the spatial processing that have been applied to the data. It is obvious that spatial filtering can be used to alter the position of the trade-off between noise level and spatial resolution. In addition to a description of the processing applied to the data, amplitude spectra of profile data were used to assist with the characterisation of the effects of processing.

Average along-line amplitude spectra were calculated for both frequency domain and time domain data examples. At any point in the spatial frequency spectrum, the observed amplitude can be attributed to a combination of geological signal, noise levels and the spatial processing that has been applied to the data. At higher spatial frequencies (ie short wavelengths), the effects of low pass filter operations often dominates. In the case of the time domain data, spatial frequencies higher than 0.6 to 0.7 Hz had been attenuated by spatial processing. At an average speed of 70 m/s, this corresponds to a wavelength of 100 to 120 metres. This is consistent with the effects of the 3 second tapered stacking filter applied to these data. In the case of the frequency domain data, spatial frequencies higher than 1.5 to 2.0 Hz had been attenuated by spatial processing. At an average speed of 33 m/s, this corresponds to a spatial wavelength of 16 to 22 metres. This is consistent with the effects of the 0.9 second tapered profile filter applied to these data.

When the residuals are computed as outlined in the previous section they show a substantial variability due to altitude variation between and along lines. As this variability is usually monitored and accommodated in processing it cannot be considered to be noise and should be removed. Figure 2 shows the residuals for the amplitude of the 385 Hz frequency data as a function of the residual computed on the radar altimeter data.

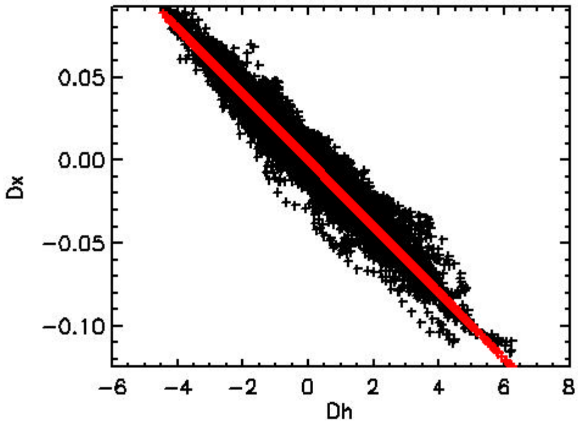

Figure 2. Scatterplot showing the dependence of the 385Hz amplitude residuals (Dx) on flying height residuals (Dh) before "height correction".

If a linear trend is removed from the data the result is much more evenly scattered. In some sense this removal amounts to a height correction of the data. This is a purely statistical "fudge" but it may not be too bad as a first approximation. All amplitudes and phases were "height corrected" in this manner.

RESULTS

The profile shown in Figure 3 is of all six amplitudes for the line with the greatest variability (line 4). They range in colour from blue for the lowest frequency to red for the highest. They have been raised to the power 10 to get back to a multiplicative factor that is an estimate for $E_{l,i}$. It can be seen that not only is the pattern the same, the plots for each frequency are also similar in amplitude. The corresponding plot for the phases shows a similar pattern although the amount of phase shift (< 1 degree) is small.

Of course, we would expect these data to be highly correlated and the correlation matrix ordered from the lowest to highest frequency for the amplitude residuals confirms this.

1.00	0.94	0.85	0.92	0.85	0.89
0.94	1.00	0.91	0.98	0.92	0.95
0.85	0.91	1.00	0.93	0.89	0.93
0.92	0.98	0.93	1.00	0.95	0.95
0.85	0.92	0.89	0.95	1.00	0.91
0.89	0.95	0.93	0.95	0.91	1.00

The Figure 3 also shows the displacement that was made to the data to bring it onto the average line (in black). Clearly some multiplicative process is changing both amplitude and phase for all channels and the error process is related to the flight parameters.

Table 1 shows the estimates for the standard deviation of the multiplicative noise when applied to the in-phase and quadrature components.

385	1581	3323	6135	25380	106140
1.4 %	1.7 %	2.0 %	1.8 %	2.2 %	2.3 %
1.8 %	2.3 %	2.8 %	2.5 %	3.3 %	5.2 %

Table 1. First row: frequency (Hz), second row: standard deviation of the in-phase residuals expressed as a percentage of the response, third row: standard deviation of the quadrature residuals.

When a similar (multiplicative noise) analysis is conducted for early time Tempest data (the first 7 channels, times as per the system described in Lane et al. (2000)). The results are somewhat different. The following points are worth noting:

- The (additive) noise measured at high altitude varies from 21 aT to 9 aT (early to late times) in the X component and 14 aT to 5 aT in the Z.
- X component residuals are much larger than for the Z component.
- The X component residuals (Figure 4).are largest at late time (σ = 2.2 % of the response) and decrease steadily as you move to earlier times (σ =1.2 %).
- The Z component residuals (Figure 5) are (usually) largest at early time (σ = 1.3 %) and show a slight decrease (σ = 1.2 %) at later times.

DISCUSSION

The Dighem results show that the multiplicative error factor appears to be roughly the same at all frequencies. Some simple modelling of halfspace responses with changing roll angle also shows that the effect is not a function of frequency. However it is also clear the effect is larger in the quadrature than the in-phase. This fact would usually indicate a residual contribution from altitude variations but the profile pattern is not the same as that of the altitude variations. More work is required to fully understand the process operating here.

In the case of the Tempest data we have a situation where the X component multiplicative factor increases with time and the Z (generally) decreases. We can use a thin sheet model to examine these effects analytically. Let the depth of the Tx image below the Rx be

$$z = (2z_a - r_z) + \frac{2t}{\mu_0 S}$$

where z_a is the aircraft altitude and r_z is the bird distance below the aircraft. When is the bird is *x* m behind the aircraft then an Rx pitch angle of β reduces the Z component response by the following factor

$$\frac{B_z(\beta)}{B_z(0)} = \frac{3xz}{x^2 - 2z^2}\sin(\beta) + \cos(\beta)$$

For large z (late time or resistive ground) the geometric term approaches zero and the factor becomes ~ cos(β). However, for the X component, the corresponding formula is

$$\frac{B_x(\beta)}{B_x(0)} = \frac{2z^2 - x^2}{3xz}\sin(\beta) + \cos(\beta)$$

where, for large z, the geometric term increases proportionally with z, making the multiplicative error strongly dependent on time and/or ground conductivity. This theory largely explains the effects we observe in the data.

It is noteworthy that this dependence of multiplicative noise on time/conductivity is a function of the asymmetric geometry of fixed wing systems time domain systems. The helicopter time domain systems are usually symmetric (*x* approximately equal to 0) causing the multiplicative errors to be constant for all times and conductivities.

CONCLUSIONS

Analysis of repeated flight lines has shown that multiplicative errors will provide the dominant source of noise for AEM surveys where the area is even moderately conductive. Errors typically have standard deviations of 2 % and can easily induce fluctuations of 10%. Moreover, because of their geometric origins, these errors are highly correlated between channels. Frequency domain HEM, Helicopter time domain systems and Z component fixed wing time domain systems produce errors that are largely unrelated to time/frequency/conductivity. This is not the case for X component data from asymmetric systems.

REFERENCES

Lane, R., Green, A., Golding, C., Owers, M., Pik, P., Plunkett, C., Sattel, D., and Thorn, B., 2000, An example of 3D conductivity mapping using the TEMPEST airborne electromagnetic system: Exploration Geophysics, 31, 162-172.

York, D., 1966, Least-square fitting of a straight line: Canadian Journal of Physics, 44, 1079-1086.

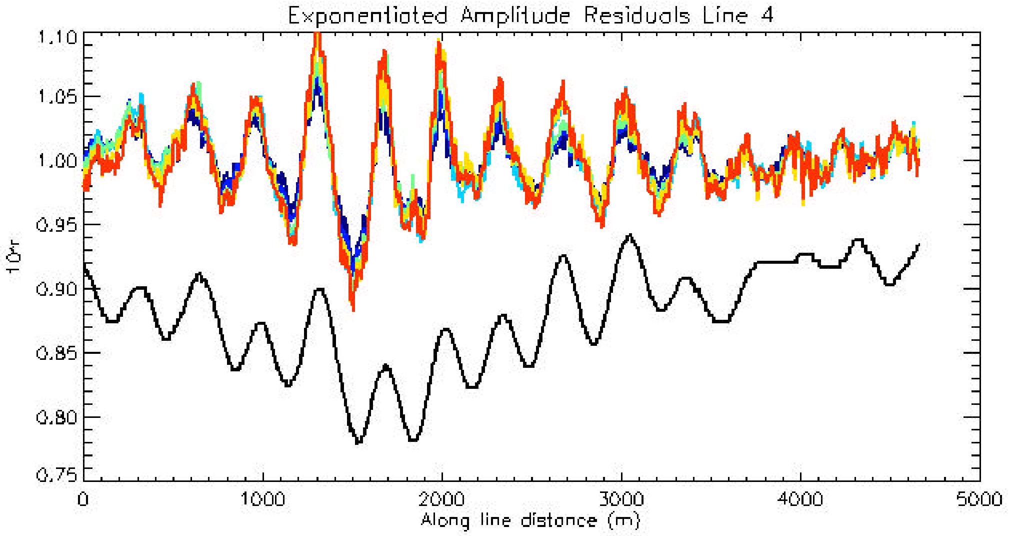

Figure 3. Dighem Line 4 amplitude residuals for all six frequencies. High frequency in red through to low in blue. The black profile is of the transverse correction, C, required to move the observation points onto the common line (rescaled, C/1000 +0.9).

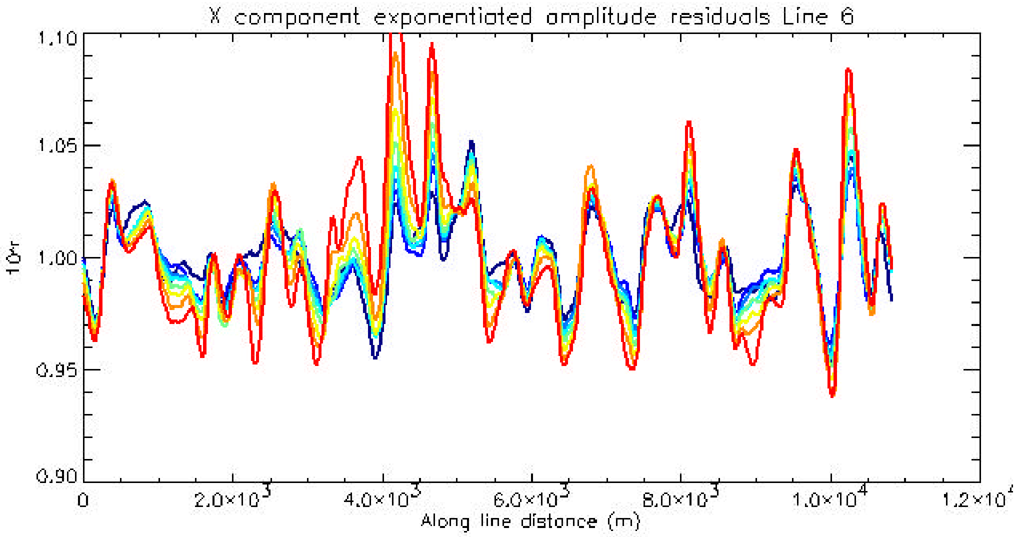

Figure 4: Tempest Line 6 X component residuals. Early times in blue through to later times in red.

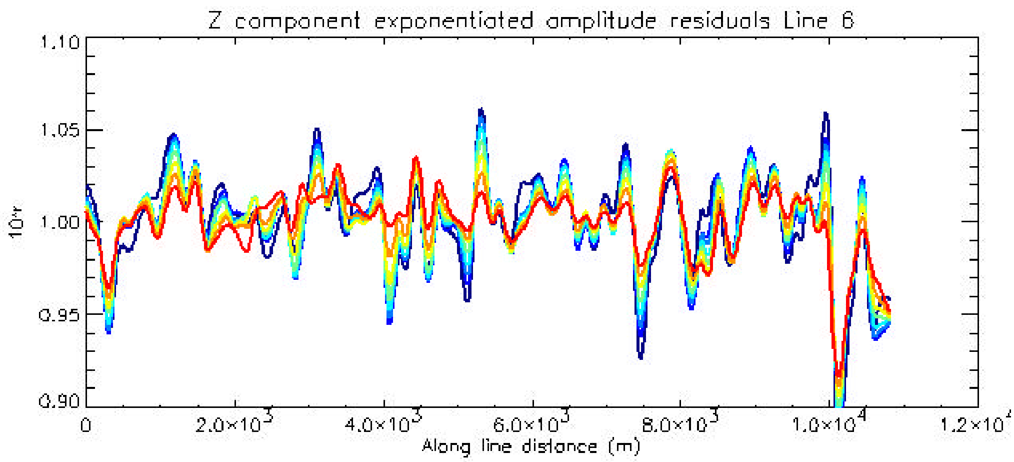

Figure 5: Tempest Line 6 Z component residuals. Early times in blue through to later times in red.