J Geophys (1983) 53: 52–58

Journal of Geophysics

The Estimation and Removal of a Linear Drift from Stacked Data

P.L. McFadden[1] and S.C. Constable

Research School of Earth Sciences, Australian National University P.O. Box 4, Canberra, ACT 2600, Australia

Abstract. In exploration systems with digital data acquisition the process of stacking is frequently performed to reduce the requirement for memory space. If, as is often the case, the observations are corrupted by a drift then this drift is not averaged out by the stacking process. The presence of such a drift will frequently cause serious errors in estimation of the signal. A robust method for least squares linear estimation and removal of the drift is presented together with an analysis of the errors involved. Examples with both synthetic and field data are presented to show the improvement in accuracy of signal recovery achieved by drift removal.

Key words: Electrical sounding – Induced polarisation – Statistics – Data processing

Introduction

In the past few years the availability of microchips and minicomputers has had a significant effect in exploration geophysics. It has now become relatively common for exploration measurements in the field to be computer controlled and for "on-site" data processing to be performed. In particular computer equipment is now being used for induced polarisation (I.P.) studies and DC resistivity soundings. Evidently signal processing cannot be applied to a true DC. However, it is possible to use a sinusoidal current with a frequency low enough that the Earth will respond in an entirely resistive manner and then apply signal processing techniques to the sinusoidal signal. I.P. work entails the use of periodic square wave signals.

Typically an electrical measurement (resistivity or I.P.) would then consist of many voltage measurements per cycle of the input current repeated over several cycles. A large memory (which is usually expensive) would be required to record each of the individual voltage measurements and to avoid this the observations are "stacked" to provide a composite or averaged set of observations over a single cycle. The simplest way to achieve this is to perform the voltage measurements at regular intervals in time and to choose this interval such that a single period of the input current (sinusoid or square wave) is an integral multiple of the measurement interval. If a fast Fourier transform (FFT) or discrete Fourier transform (DFT) is then applied to the stacked observations the signal frequency is precisely equal to the fundamental frequency of the transform and so there is no side-lobe leakage or phase shifting of the fundamental frequency due to the finite record length.

Unfortunately electrical measurements are frequently plagued by potential electrode drift or a drift in telluric potentials. As will be shown, a linear drift is not averaged out by the process of stacking many cycles back to a single cycle of the signal and, from the point of view of an FFT or DFT, the drift appears in the stacked waveform as a sawtooth wave with fundamental frequency equal to the signal frequency. Consequently the presence of drift can lead to grossly inaccurate results and so it is vital to remove the drift from the stacked waveform before performing an FFT.

In the process of stacking information is lost and it becomes very difficult to obtain a reliable estimate of the drift for subsequent removal. A rough estimate can be obtained by taking the difference between the first and last points of the stacked waveform. However, this actually makes use of very little information and so is subject to large amount of random error. If a simple least squares linear regression is performed on the stacked waveform this will lead to an *incorrect* estimate of the drift. This is easily seen from the fact that a least squares linear regression applied to one cycle of the function $(A\cos\theta + B\sin\theta)$ will give an estimated linear drift of $(-6B/\pi)$ per cycle, when in fact none exists.

The purpose of this paper is to suggest a simple modification to the real-time stacking process so that relevant information is not lost, thus allowing optimal estimation and removal of a linear drift. An analysis of the statistical properties of this process is presented so that a better understanding of the results can be achieved. Finally some examples of synthetic and actual field data are presented to show how effective the method is.

Formulation of the Problem

The signal may be considered merely as some periodic function $F(\theta)$ such that

$$F(\theta + 2\pi) = F(\theta) \tag{1}$$

and the measurements of $F(\theta)$ are performed at n equally spaced positions per period for N consecutive per-

[1] *Present address:* Scripps Institution of Oceanography UCSD, la Jolla, Ca 92093, USA

iods. The period number will always be indexed as i ($i=1$ to N) and the position in the period as j ($j=1$ to n) so that

$$F_{ji}=F(\theta_j+2\pi(i-1))=F(\theta_j),$$
$$\theta_j=\frac{2\pi j}{n}. \tag{2}$$

Evidently there will always be random noise and so any direct measurement of F_{ji} would give an observation f_{ji},

$$f_{ji}=F_{ji}+\varepsilon, \tag{3}$$

where ε has expectation zero and variance σ_f^2. For the present it is assumed that this variance is constant for all i and j. We now assume that any measurement of F_{ji} is corrupted by a linear drift of D per period and a DC offset C. There will also be random noise in the drift and so any direct measurement (i.e. no signal) of D would give an observation d,

$$d\left(\frac{\theta_j}{2\pi}+i-1\right)=D\left(\frac{\theta_j}{2\pi}+i-1\right)+\delta, \tag{4}$$

where δ has expectation zero and variance σ_d^2. Again we assume for the present that this variance is constant.

With this model a particular observation, y_{ji}, of the voltage will be given by

$$y_{ji}=f_{ji}+d\left(\frac{\theta_j}{2\pi}+i-1\right)+C$$
$$=f_{ji}+d\left(\frac{j}{n}+i-1\right)+C. \tag{5}$$

If these observations are stacked in real time to obtain a composite waveform over a single period the resulting observation is obtained by summing over i from 1 to N as

$$\sum_{i=1}^{N} y_{ji}=\sum_{i=1}^{N} f_{ji}+\frac{Njd}{n}+\tfrac{1}{2}N(N-1)\,d+NC, \tag{6}$$

giving the average values, $\bar{y}_j$ and $\bar{f}_j$, for each stack as

$$\bar{y}_j=\bar{f}_j+\frac{jd}{n}+\tfrac{1}{2}(N-1)\,d+C. \tag{7}$$

These n observations $\bar{y}_j$ are then the only observations available, as the individual y_{ji} are lost in the stacking process, and so the signal $F(\theta)$ must be estimated from the n values of $\bar{f}_j$ given by

$$\bar{f}_j=\bar{y}_j-\frac{jd}{n}-\tfrac{1}{2}(N-1)\,d-C. \tag{8}$$

The term $\frac{1}{2}(N-1)\,d$ may be considered merely as an additional DC offset but the drift still enters through the term (jd/n). Hence the drift must be estimated and removed.

Estimation and Removal of the Drift

In order to simplify the notation the subscript j will now be dropped from double subscripts so that $y_i\equiv y_{ji}$, $f_i\equiv f_{ji}$ and the summation signs are for i from 1 to N unless otherwise specified. Considering a particular stack j an unbiased estimate for the variance in the f_i is given by s_j^2 where

$$(N-1)\,s_j^2=\Sigma(f_i-\bar{f})^2 \tag{9}$$

where $\bar{f}\equiv\bar{f}_j$. Any drift unaccounted for will inflate this sum and so the drift D may be estimated as that value of d in Eq. (8) which minimises s_j^2. Substituting from Eqs. (5) and (8) into (9) gives

$$(N-1)\,s_j^2=\tfrac{1}{12}(N^3-N)\,d^2+[(N+1)\,\Sigma y_i-2\Sigma i\,y_i]\,d$$
$$+\Sigma(y_i-\bar{y})^2 \tag{10}$$

where $\bar{y}\equiv\bar{y}_j$. Differentiation with respect to d then gives

$$(N-1)\frac{\partial s_j^2}{\partial d}=\tfrac{1}{6}(N^3-N)\,d+(N+1)\,\Sigma y_i-2\Sigma i\,y_i \tag{11}$$

and so $\hat{d}_j$, the value of d for which the above differential is zero (and therefore the value of d which minimises s_j^2), is given by

$$\hat{d}_j=\frac{2\Sigma i\,y_i-(N+1)\,\Sigma y_i}{\frac{1}{6}(N^3-N)}. \tag{12}$$

Each of the n stacks will lead to an estimate $\hat{d}_j$ and so an overall estimate, $\bar{d}$, of D will be given by the mean value as

$$\bar{d}=\frac{1}{n}\sum_{j=1}^{n} d_j. \tag{13}$$

In summary, in order to remove a linear drift it is necessary to accumulate for each stack j not only the sums of the measurements, Σy_i, but also the sums of the measurements multiplied by the stack number, $\Sigma i\,y_i$. These are then substituted into (12) to give n independent estimates $\hat{d}_j$ of the drift per cycle, D. The overall estimate, $\bar{d}$, of D is then taken as the average of the n values $\hat{d}_j$ (Eq. 13). This value may then be substituted for d in Eq. (8) to remove the effect of the drift. Further, if it is known that the mean of $F(\theta)$ over a cycle is zero then, without loss of generality, C may be chosen as that value which makes the sum of the n values $\bar{f}_j$ (corrected for drift) equal to zero.

Properties of the Drift Estimator

The estimate $\hat{d}_j$ is a linear combination of the observations y_i and consequently it has several useful properties. Using Eqs. (3) and (4) Eq. (5) may be rewritten as

$$y_i=F_i+\varepsilon+D\left(\frac{j}{n}+i-1\right)+\delta+C$$
$$=(F_i+C)+Y_i \tag{14}$$

where Y_i is merely an observation of the drift corrupted by both the noise (δ) in the drift and the noise (ε) in the signal. Substitution of the y_i values in this form into Eq. (12) gives

$$\hat{d}_j=\frac{2\Sigma i\,Y_i-(N+1)\Sigma Y_i}{\frac{1}{6}(N^3-N)}. \tag{15}$$

If we now let $x_i \equiv i$ and recognise that

$$(N+1)=\frac{2}{N}\Sigma i\equiv\frac{2}{N}\Sigma x_i=2\bar{x} \tag{16}$$

and

$$\tfrac{1}{6}(N^3-N)=2\Sigma\left(i-\frac{N+1}{2}\right)^2\equiv 2\Sigma(x_i-\bar{x})^2 \tag{17}$$

then

$$\hat{d}_j=\frac{\Sigma(x_i-\bar{x})Y_i}{\Sigma(x_i-\bar{x})^2}=\frac{\Sigma(Y_i-\bar{Y})x_i}{\Sigma(x_i-\bar{x})^2}. \tag{18}$$

Hence $\hat{d}_j$ is merely the slope of a least squares linear regression of the drift. Expansion of Eq. (15) with the error terms ε and δ shows that the expectation of $\hat{d}_j$ is D and the variance, σ_D^2, of $\hat{d}_j$ is given by

$$\sigma_D^2=\frac{12}{N^3-N}(\sigma_f^2+\sigma_d^2)$$
$$=\frac{12}{N^3-N}\sigma_t^2 \tag{19}$$

where σ_t^2 is the overall variance of the noise. Each of the $\hat{d}_j$ is an independent estimate of D and so an unbiased estimate for the variance σ_D^2 is given by s_D^2 where

$$s_D^2=\frac{1}{n-1}\sum_{j=1}^{n}(\hat{d}_j-\bar{d})^2. \tag{20}$$

Evidently an unbiased estimate for the variance in $\bar{d}$ is given by (s_D^2/n).

If the errors ε and δ are normally distributed then the $\hat{d}_j$ are normally distributed with mean D and variance σ_D^2. Furthermore the estimate s_D^2 will have a chi-square distribution with $(n-1)$ degrees of freedom as

$$\frac{(n-1)s_D^2}{\sigma_D^2}\sim X^2_{(n-1)}. \tag{21}$$

Here the symbol "$\sim$" is to be read as "is distributed as".

Finally, the fact that the $\hat{d}_j$ are least squares linear regressions means the overall estimate $\bar{d}$ is fairly robust to deviations from a truly linear drift.

Applications when $F(\theta)$ is a Single Sinusoid

If $F(\theta)$ is a single frequency sinusoid then

$$F(\theta_j)=A\cos\left(\frac{2\pi j}{n}\right)+B\sin\left(\frac{2\pi j}{n}\right) \tag{22}$$

and A and B are given by

$$A=\frac{2}{n}\sum_{j=1}^{n}F(\theta_j)\cos\left(\frac{2\pi j}{n}\right)$$
$$B=\frac{2}{n}\sum_{j=1}^{n}F(\theta_j)\sin\left(\frac{2\pi j}{n}\right). \tag{23}$$

Having corrected the stacked data for the drift the values $\bar{f}_j$ are estimates of $F(\theta_j)$ and so estimates of A and B may be obtained. In terms of understanding the estimates so obtained it is important to investigate the statistical properties of these estimates.

If we define $\bar{\delta}$

$$\bar{\delta}=D-\bar{d} \tag{24}$$

then $\bar{\delta}$ is the error in estimation of the drift and so after correcting the stacked results for the drift there is still a drift of $\bar{\delta}$ per cycle left. We now consider a particular stack j again. In Eq. (10) the term $\Sigma(y_i-\bar{y})^2$ may be expanded as

$$\Sigma(y_i-\bar{y})^2=\Sigma y_i^2-\frac{1}{N}(\Sigma y_i)^2 \tag{25}$$

and so if the sum of the squares of the observations is also kept the value of $\Sigma(f_i-\bar{f})^2$ for each stack may be determined from Eq. (10) after substituting $\bar{d}$ for d. Using Eq. (10) in this manner it follows that

$$\left\langle\left(\frac{\Sigma(f_i-\bar{f})^2}{N-1}\right)\right\rangle=\langle s_j^2\rangle=\sigma_t^2+\tfrac{1}{12}(N^2+N)\langle\bar{\delta}^2\rangle \tag{26}$$

where the notation $\langle z\rangle$ indicates the expectation of z. In the general sense of repeating the whole experiment many times, $\bar{\delta}$ is a random variable and $\langle\bar{\delta}^2\rangle$ is simply the variance of $\bar{d}$, i.e. (σ_D^2/n). Substituting from Eq. (19) then gives

$$\langle s_j^2\rangle=\left(1+\frac{1}{n(N-1)}\right)\sigma_t^2 \tag{27}$$

and so even for fairly small N and n it may be considered that $\langle s_j^2\rangle=\sigma_t^2$. However, for the case of a single experiment (the usual case), $\bar{\delta}$ is not a random variable but a constant and so $\langle\bar{\delta}^2\rangle=\bar{\delta}^2$. From the argument leading to Eq. (27) it is evident that $\frac{1}{12}(N^2+N)\bar{\delta}^2$ will be small compared to σ_t^2 and so for a single experiment it may also be considered that $\langle s_j^2\rangle=\sigma_t^2$.

Using the $\bar{f}_j$ as estimates of $F(\theta_j)$, estimates a and b for A and B are obtained as

$$a=\frac{2}{n}\sum_{j=1}^{n}\bar{f}_j\cos\left(\frac{2\pi j}{n}\right)$$
$$b=\frac{2}{n}\sum_{j=1}^{n}\bar{f}_j\sin\left(\frac{2\pi j}{n}\right). \tag{28}$$

The estimate, s_a^2, for the variance, σ_a^2, of a is given by

$$s_a^2=\frac{4}{n^2N}\sum_{j=1}^{n}s_j^2\cos^2\left(\frac{2\pi j}{n}\right) \tag{29}$$

since (s_j^2/N) is the estimate of the variance in $\bar{f}_j$. Furthermore

$$\langle s_a^2\rangle=\sigma_a^2=\frac{4}{nN}\langle s_j^2\rangle\frac{1}{n}\sum_{j=1}^{n}\cos^2\left(\frac{2\pi j}{n}\right)$$
$$=\frac{2}{nN}\sigma_t^2. \tag{30}$$

Similarly

$$s_b^2=\frac{4}{n^2N}\sum_{j=1}^{n}s_j^2\sin^2\left(\frac{2\pi j}{n}\right) \tag{31}$$

and

$$\langle s_b^2\rangle=\sigma_b^2=\frac{2}{nN}\sigma_t^2=\sigma_a^2. \tag{32}$$

The residual drift $\bar{\delta}$ per cycle affects the estimates a and b so that

$$\langle a \rangle = A; \quad \langle b \rangle = B - \frac{1}{\pi}\bar{\delta}. \tag{33}$$

Consequently a is an unbiased estimate of A but b is not quite an unbiased estimate of B. However, the bias in b is small enough that in practice it may be ignored.

Invoking the central limit theorem the estimates a and b will be normally distributed independent of the original distributions of the noise. Consequently it is a simple matter to obtain error limits for A and B.

Testing for a Non-linear Drift

If the original noise in the signal and the drift is normally distributed then the estimates s_j^2 will each have chi-square distributions with $(N-1)$ degrees of freedom. The average of the n values of s_j^2 will also have a chi-square distribution but with $n(N-1)$ degrees of freedom. However, it is s_a^2 and s_b^2 which are calculated rather than the average of the s_j^2 values. Consequently it is more useful to determine the distribution of these estimates. Reference is made here to s_a^2 but because of the equivalence of s_a^2 and s_b^2 it is immaterial whether we use s_a^2, s_b^2 or $\frac{1}{2}(s_a^2+s_b^2)$.

In determining s_a^2 the individual values of s_j^2 are multiplied by different factors and so the distribution of s_a^2 cannot be exactly chi-square. However, it will be approximately chi-square distributed and because $n(N-1)$ will represent a large number of degrees of freedom this approximation will be very good. Hence we can choose as the approximating chi-square distribution that which has the same mean and variance as the statistic s_a^2. Thus, as a very good approximation,

$$s_a^2 \sim \frac{3}{n^2}\frac{\sigma_t^2}{N(N-1)}X^2_{\frac{2}{3}n(N-1)}. \tag{34}$$

Combining Eqs. (19) and (21)

$$s_D^2 \sim \frac{12\sigma_t^2}{(n-1)(N^3-N)}X^2_{(n-1)}. \tag{35}$$

It then follows from Eqs. (34) and (35) that

$$\frac{s_a^2}{s_D^2}\frac{6n}{N^2-1} \equiv G \sim F[\frac{2}{3}n(N-1),(n-1)], \tag{36}$$

where $F[m_1, m_2]$ is the F distribution with m_1 and m_2 degrees of freedom. The statistic G of Eq. (36) is merely the ratio of σ_t^2 estimated from the scatter in the f_i values to σ_t^2 estimated from the scatter in the individual drift estimates $\hat{d}_j$. Evidently in practice m_1 must be chosen as the integer nearest $2n(N-1)/3$.

If the drift was in fact non-linear then systematic errors due to misfit of the model will cause an apparent inflation of σ_d^2 and so, from Eq. (19), will cause an apparent inflation of σ_t^2 as estimated from scatter in the individual drift estimates. This apparent inflation will also appear in the s_j^2 through σ_t^2 in Eq. (26). However, the second term of Eq. (26), $(N^2+N)\ \langle\bar{\delta}^2\rangle/12$, will be replaced by the term $(\Sigma(i-\frac{1}{2}[N+1])^2\,\delta_i^2)/(N-1)$ where the δ_i represent the systematic errors due to misfit of the model. Consequently the second term of Eq. (26) would no longer be small with respect to the first term but would instead be of the same order. Hence if σ_t^2 is estimated from the scatter in the f_j values then this estimate would be inflated substantially more than the estimate of σ_t^2 determined from scatter in the individual drift estimates. Under these conditions the statistic G of Eq. (36) would not have an F-distribution and we would expect an observed value of the statistic to be substantially greater than unity. Thus if the observed value of G exceeds the critical value of the F-distribution then the hypothesis of a linear drift may be rejected.

It is important to recognise that the statistical ability to reject the hypothesis of a linear drift does not invalidate the procedure suggested here for removal of drift. If the random noise in an experiment is very small then a very small amount of non-linearity in the drift will cause rejection of the linear drift hypothesis. Conversely, if the random noise in an experiment is large then a moderate degree of non-linearity in the drift will not cause rejection of the linear drift hypothesis. Equation (28) for estimating A and B is merely the fundamental frequency term of a DFT applied to the data and, as noted previously, in a stacked waveform a linear drift contributes to the power observed in this fundamental frequency. This is a consequence of the stacking process and holds true for a non-linear drift as well. Estimation and removal of the drift as if it were a linear drift will effectively remove most of the apparent power in the fundamental frequency consequent upon stacking a non-linear drift. However, it should be clear from the preceding argument that the estimate of the variance in the result will be artifically inflated.

Examples of Drift Removal

Synthetic data

In order to show the effectiveness of the drift removal process, synthetic data were generated from

$$y_{ji} = A\cos\theta_j + B\sin\theta_j + C + D\left(\frac{\theta_j}{2\pi}+i-1\right) + Q\left(\frac{\theta_j}{2\pi}+i-1\right)^2 + \text{Gaussian noise (variance } \sigma_t^2). \tag{37}$$

Here Q is a quadratic drift per cycle to simulate a non-linear drift and the other terms are as previously defined.

In the first example $A=0$ and there is only a linear drift, with the drift per cycle being twice the signal amplitude. Figure 1 shows the estimated signal and data both before and after estimation and removal of the linear drift. The details of the parameters are given in Table 1. In the second example there is still only a linear drift (but twice that of the first example) and A is now non-zero. The fits are shown in Fig. 2 and the details are given in Table 2. In both examples it may be seen that the presence of a linear drift affects the estimation of B far more than the estimation of A, as is to be expected from Eq. (33) with $\bar{\delta}$ replaced by D. In neither example can the hypothesis of a linear drift be

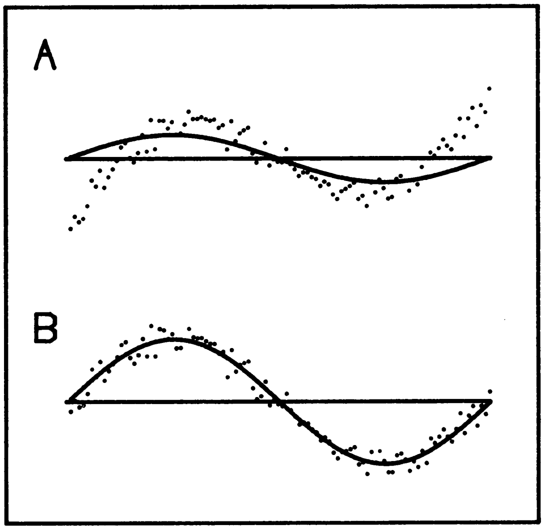

Fig. 1a and b. Synthetic data with a linear drift. **a** before and **b** after estimation and removal of drift. See Table 1 for details

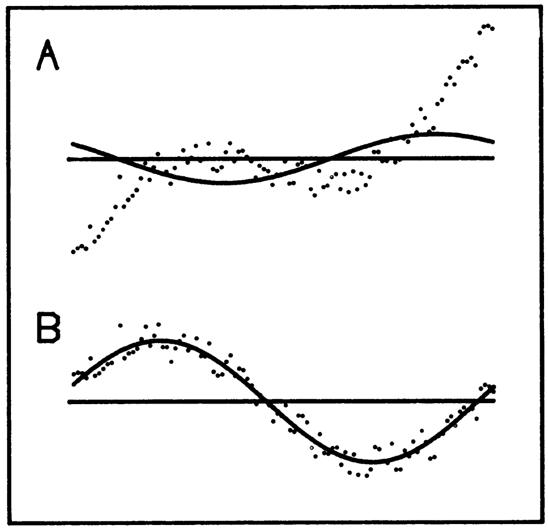

Fig. 2a and b. Synthetic data with a linear drift. **a** before and **b** after estimation and removal of drift. See Table 2 for details

Table 1. Synthetic data with a linear drift. See also Fig. 1

Actual parameters	Estimated parameters	
	Before drift removal	After drift removal
$A = 0.0$	$a = 0.0001$ $s_a = 0.2720$	$a = -0.020$ $s_a = 0.022$
$B = 1.0$	$b = 0.39$ $s_b = 0.27$	$b = 1.028$ $s_b = 0.022$
$C = 2.0$	$c = 12.03$	$c = 2.03$
$D = 2.0$		$\bar{d} = 1.998$ $s_D = 0.054$
$Q = 0.0$		
$\sigma_t = 0.5$		$s_1 = 0.489$ $s_2 = 0.490$
$n = 100$		$G = 1.06$
$N = 10$		

Notes

1. s_1 is an estimate of σ_t obtained by substituting s_D^2 for σ_D^2 in Eq. (19)
2. $s_2^2 = \frac{1}{2} n N s_a^2$ and so s_2 is also an estimate of σ_t (see Eq. 30)
3. Critical value of G at the 95% level of confidence = 1.29

Table 2. Synthetic data with a linear drift. See also Fig. 2

Actual parameters	Estimated parameters	
	Before drift removal	After drift removal
$A = 0.25$	$a = 0.268$ $s_a = 0.543$	$a = 0.228$ $s_a = 0.023$
$B = 1.00$	$b = -0.297$ $s_n = 0.544$	$b = 0.980$ $s_b = 0.023$
$C = 1.00$	$c = 21.05$	$c = 0.973$
$D = 4.00$		$\bar{d} = 4.012$ $s_D = 0.057$
$Q = 0.00$		
$\sigma_t = 0.50$		$s_1 = 0.518$ $s_2 = 0.514$
$n = 100$		
$N = 10$		$G = 0.992$

Notes

1. s_1 and s_2 are as defined in Table 1
2. Critical value of G at the 95% level of confidence = 1.29

rejected (as expected). Both of these examples show that the presence of a linear drift can produce serious errors in signal estimation and that the drift removal process suggested here is very effective in such instances.

In the third example a fairly severe quadratic drift has been used. The fits are shown in Fig. 3 and the details are given in Table 3.

Since a stacked quadratic has even function components the estimation of A is now also affected seriously. As expected, the hypothesis of a linear drift may be rejected and this has some consequences which are of interest.

First of all, the DC component, C, is incorrectly estimated (even after removal of a least squares linear drift), but this is of little importance unless C is itself a desired parameter. Secondly, the noise, s_2, estimated from s_b is much larger than the estimate s_1, from s_D, and both of these are much larger than the actual noise, σ_t. Thus, as expected from the theory, if the hypothesis of a linear drift can be rejected then the calculated errors on the parameter estimates are in fact too large. However, it may be seen from this example that the drift removal process is very robust, the overall amplitude of the signal being estimated to within 2% and the phase to within 0.6° (without any drift removal the amplitude estimate is only 35% of the correct amplitude and the estimated phase would be in error by 39°).

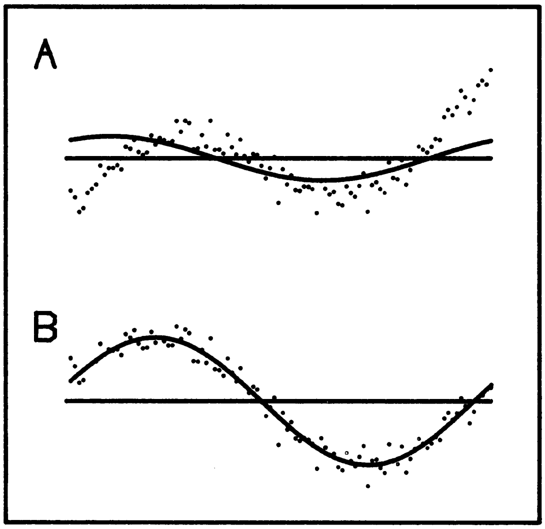

Fig. 3a and b. Synthetic data with a quadratic drift. **a** before and **b** after linear estimation and removal of drift. See Table 3 for details

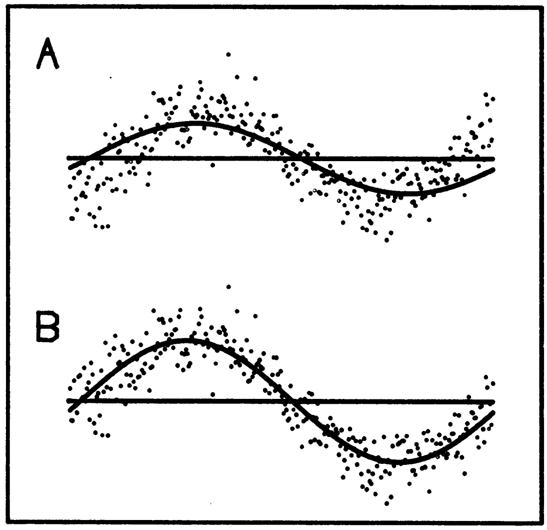

Fig. 4a and b. Field data. **a** before and **b** after linear estimation and removal of drift. See Table 4 for details

Table 3. Synthetic data with a quadratic drift. See also Fig. 3

Actual parameters	Estimated parameters	
	Before drift removal	After drift removal
$A=0.25$	$a=0.290$	$a=\ 0.265$
	$s_a=0.351$	$s_a=\ 0.091$
$B=1.00$	$b=0.219$	$b=\ 1.017$
	$s_b=0.351$	$s_b=\ 0.091$
$C=1.00$	$c=9.304$	$c=-3.230$
$D=0.00$		$\bar{d}=\ 2.504$
		$s_D=\ 0.159$
$Q=0.25$		
$\sigma=0.50$		$s_1=\ 1.444$
		$s_2=\ 2.035$
$n=100$		
$N=10$		$G=\ 1.979$

Notes

1. s_1 and s_2 are as defined in Table 1
2. Critical value of G at the 95% level of confidence $=1.29$
3. If $c'+\bar{d}'i$ is a least squares linear fit to the quadratic drift (before stacking), then $c'=-3.179$ and $\bar{d}'=2.503$, in excellent agreement with the estimate from the stacked data

Table 4. Field data. See also Fig. 4

Before drift removal	After drift removal
$a=-0.316$ mV	$a=-0.324$ mV
$s_a=\ 0.438$ mV	$s_a=\ 76.3$ μV
$b=\ 1.01$ mV	$b=\ 1.79$ mV
$s_b=\ 0.336$ mV	$s_b=\ 80.6$ μV
$c=-0.930$ V	$c=-0.968$ V
	$\bar{d}=\ 2.48$ mV
	$s_D=\ 77.2$ μV
	$s_1=\ 3.66$ mV
	$s_2=\ 5.41$ mV
	$G=\ 2.18$

Notes

1. s_1 and s_2 are as defined in Table 1
2. Critical value of G at the 95% level of confidence $=1.16$
3. $n=300$ and $N=30$

Field Data

It is to be expected that in the real world a truly linear drift would be a rare phenomenon. However, as has already been shown, the drift removal process is robust to non-linearities and therefore provides a very useful tool in the analysis of real data.

Figures 4 and 5 show two examples of field data obtained using a computer controlled electrical sounding system. The details are given in Tables 4 and 5. The input signal is a very low frequency, in-phase sinusoidal current. The recovered voltage signal is very noisy, phase shifted and corrupted by a large drift. In both cases the hypothesis of a linear drift may be rejected but the great improvement in signal recovery after drift removal is immediately apparent.

The final example of real data is quite interesting, the fits being shown in Fig. 6 and the details given in Table 6. Here the noise is very small, a drift is readily apparent and there is a distinct "kink" in the data. Because the noise is small this "kink" appears as a strong non-linearity in the drift and the hypothesis of a linear drift is rejected (very strongly so in this case). Again, the improvement in signal recovery after drift removal is apparent.

Conclusions

A method for removal of a linear drift from stacked data has been presented. In addition it has been shown that the method is robust to non-linearities (because it is a least squares fit) and thus is a useful tool in the analysis of real data.

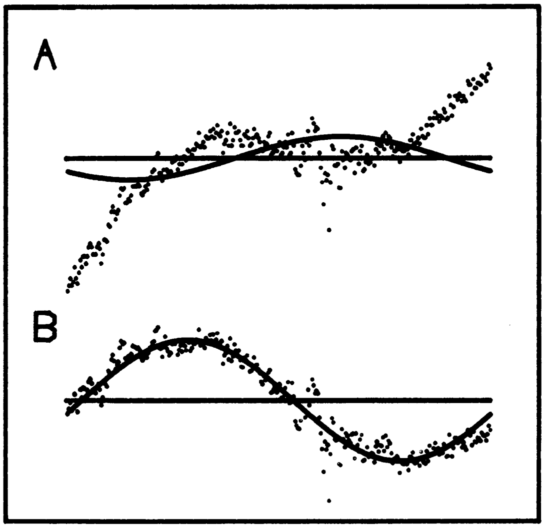

Fig. 5 a and b. Field data. **a** before and **b** after linear estimation and removal of drift. See Table 5 for details

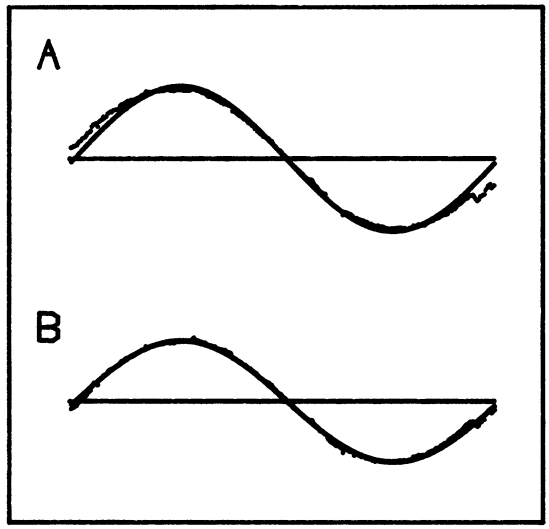

Fig. 6 a and b. Field data. **a** before and **b** after linear estimation and removal of drift. See Table 6 for details

Table 5. Field data. See also Fig. 5

Before drift removal	After drift removal
$a=-0.429$ mV	$a=-0.456$ mV
$s_a=$ 0.688 mV	$s_a=$ 6.22 μV
$b=-0.575$ mV	$b=$ 1.96 mV
$s_a=$ 0.628 mV	$s_b=$ 63.0 μV
$c=-2.07$ V	$\bar{c}=-2.11$ V
	$\bar{d}=$ 7.96 mV
	$s_D=$ 0.205 mV
	$s_1=$ 1.86 mV
	$s_2=$ 2.44 mV
	$G=$ 1.72

Notes

1. s_1 and s_1 are as defined in Table 1
2. Critical value of G at the 95% level of confidence = 1.17
3. $n=300$ and $N=10$

Table 6. Field data. See also Fig. 6

Before drift removal	After drift removal
$a=-0.549$ mV	$a=-0.535$ mV
$s_a=$ 0.636 mV	$s_a=$ 56.6 μV
$b=$ 8.02 mV	$b=$ 6.71 mV
$s_b=$ 0.327 mV	$s_b=$ 55.6 μV
$c=-0.306$ V	$\bar{c}=-0.285$ V
	$\bar{d}=-4.12$ mV
	$s_D=$ 74.0 μV
	$s_1=$ 0.672 mV
	$s_2=$ 2.15 mV
	$G=$ 10.3

Notes

1. s_1 and s_2 are as defined in Table 1
2. Critical value of G at the 95% level of confidence = 1.17
3. $n=300$ and $N=10$

The process of stacking reduces the required memory space by keeping only the sums of the observations at each stack point rather than retaining all of the individual observations. In order to estimate and then remove a drift the sums of (observations multiplied by cycle number) must also be kept at each stack point. The drift may then be estimated using Eqs. (12) and (13) and then removed by using Eq. (8). Although this doubles the memory required for the same number of stack points the improvement in accuracy of signal recovery fully justifies the process.

In order to perform a statistical analysis of the signal estimation the sums of the squares of the observations at each stack point must also be retained. Again this increases the memory required for a given number of stacks and in many applications it may be felt that this increase in memory requirement is not justified.

Naturally it is possible to obtain a least squares quadratic fit, from stacked data. However, this would require that the sums of (observations multiplied by the square of the cycle number) also be kept for each stack point. Since the simpler process suggested here is robust to non-linearities the increased requirement in memory size is probably not justified.

Finally, it is important to note that the statistical analysis presented here estimates only the precision of a particular measurement. Repeated measurements may show that larger errors are indicated due to a non Gaussian and/or a non stationary noise process. Above all, it must be realised that "geological noise" (the consequence of real geology being more complicated than models used to describe it) cannot be estimated by the processes described here.

Received July 6, 1982; Revised Version January 18, 1983
Accepted January 19, 1983