Determination of electrical signal masked by random noise using «super-averaged» functions

Luigi Alfano and Mauro Giudici
Università degli Studi di Milano, Dipartimento di Scienze della Terra, Sezione di Geofisica, Milano, Italy

Abstract
Electrical and electromagnetic methods often require the determination of the amplitude of a periodic signal, produced by controlled sources and masked by natural and artificial noise. Since noise is usually random and uncorrelated, several techniques based on a stacking procedure and spectral decomposition are applied. In this paper we revise some of these procedures based on Fourier analysis. We propose a technique which uses «super-averaged» functions obtained from the average of the results of the Fourier analysis from the whole data set.

Key words *electrical methods – electromagnetic methods – data acquisition – data processing – signal-to-noise ratio*

1. Introduction

Direct current electrical soundings for deep studies often require great distances between current and potentiometric dipoles. Since the electric field decreases with distance, the signal to be measured at the potentiometric dipole could be very low and masked by the electrical noises produced by natural and artificial sources.

Some methods have been proposed to evaluate the difference of electric potential at the potentiometric dipole, even when the Signal-to-Noise Ratio (SNR) is very low (Loddo and Patella, 1977; Alfano *et al.*, 1982; Ciminale and Patella, 1982; Lapenna *et al.*, 1987, 1994). In fact, the useful signal for electrical and electromagnetic prospecting is usually periodic and generated by a controlled source; it is masked by non-periodic disturbances which might be relatively large and produced by casual phenomena (noise).

In this paper we describe the techniques of stacking and Fourier analysis and a technique that we have used with success in very unfavourable situations. In the second section we recall the basic principles of the stacking procedure and Fourier analysis and we show the link between them. In the third section we analyse some characteristics of a standard «time dependent» Fourier analysis. In the fourth section we introduce the «time independent» Fourier analysis, which is the basis of our technique, described in the fifth section and applied to synthetic and real data in the sixth section.

Before entering into the core of the paper we fix some notation. We denote with V_h the elements of a numerical series, which represent, for instance, the voltages measured at constant time intervals during an electrical or electromagnetic survey. We assume that the signal, generated with controlled sources, is periodic

Mailing address: Prof. Mauro Giudici, Università degli Studi di Milano, Dipartimento di Scienze della Terra, Sezione di Geofisica, Via L. Cicognara 7, 20129 Milano, Italy; e.mail: Mauro.Giudici@unimi.it

with a period T which is an integer multiple of the sampling interval, Δ, so that $T = J\Delta$. Therefore we can indicate the signal period with J, which means that we are using the sampling rate as the measurement unit for time.

In order to describe some properties of stacking and Fourier analysis, it is useful to define the «operating period» as a collection of J successive elements of the data series; we number the operating periods with the index n and the elements of the series within an operating period with the index i. Then we can write $V_h = V_{(n-1)J+i}$. It is convenient to arrange the data series in a table, where each row includes the data belonging to one operating period

$$\begin{matrix} V_1 & V_2 & \cdots & V_J \\ V_{J+1} & V_{J+2} & \cdots & V_{2J} \\ \vdots & \vdots & \vdots & \vdots \\ V_{(Q-1)J+1} & V_{(Q-1)J+2} & \cdots & V_{QJ} \end{matrix} \qquad (1.1)$$

The rows of (1.1) correspond to the Q operating periods contained in the recorded numerical series.

Each element of the data series is given by the sum of the signal contribution, S_i, which corresponds to the artificial voltage caused by electrical currents injected into the ground, a random noise with zero mean, $N_{(n-1)J+i}$, and more regular disturbances, whose average value is not constant during data acquisition, $\Phi_{(n-1)J+i}$, so that

$$V_{(n-1)J+i} = S_i + N_{(n-1)J+i} + \Phi_{(n-1)J+i}. \qquad (1.2)$$

We are interested in the sinusoidal component of the signal with period $T = J\Delta$ and therefore we assume

$$S_i = S \sin(2\pi i / J + \varphi), \qquad (1.3)$$

where we have supposed that the phase φ of the signal is unknown, which is the case in the field when it is not possible to synchronise source and receiver.

We assume that high-frequency components have been removed with anti-alias filters before the application of the techniques for signal extraction.

2. The stacking procedure and the Fourier analysis

The stacking procedure and Fourier analysis are well known to geophysicists and are described in any textbook of applied geophysics (see, *e.g.*, Telford *et al.*, 1990). We briefly recall their characteristics, for the specific application to the determination of the periodic signal in electrical and electromagnetic prospecting.

2.1. *The stacking procedure*

The stacking procedure permits an evaluation of the signal component by computing

$$U_i = \frac{1}{Q}\sum_{n=1}^{Q} V_{(n-1)J+i}, \quad i = 1, \ldots, J. \qquad (2.1)$$

The quantities U_i given by (2.1) are the averages of the columns of (1.1).

If the data series contains rigorously non-periodic random noise only, namely if $V_{(n-1)J+i} = N_{(n-1)J+i}$, the data of each column are generally different from each other and they are distributed in a random way. Now, according to the probability principles, the limits of the mean values U_i must tend to zero for sufficiently great values of Q, the number of analysed periods, namely $\lim_{Q\to\infty} U_i = 0$.

In practice, each element of the data series is given by (1.2), so that if we could measure infinitely many operating periods,

$$\lim_{Q\to\infty} U_i = S_i + \langle \Phi_i \rangle, \qquad (2.2)$$

where $\langle \Phi_i \rangle$ denotes the ensemble average of $\Phi_{(n-1)J+i}$, when $n = 1,\ldots,\infty$. If the shape of the signal is known and if $\langle \Phi_i \rangle$ is sufficiently smooth within the operating period, *i.e.* as a function of i, the effect of the second term of eq. (2.1) may be neutralised and the stacking procedure can be applied to evaluate the signal. Actually, it is evident from (2.2) that the stacking procedure can be effective if a sufficient number of operating periods is available; namely, if the recording time has been sufficiently long.

2.2. *The Fourier transform*

We refer the reader to specific books (*e.g.*, Oppenheim and Schafer, 1989) for complete descriptions of Fourier transforms of discrete-time signals. Here we focus on the single component of the Fourier transform corresponding to the frequency $1/J$.

Let us consider the two sinusoidal functions with period J

$$C_i^{(s)} = \sin(2\pi i / J) \text{ and } C_i^{(c)} = \cos(2\pi i / J). \quad (2.3)$$

With these functions we can introduce the Fourier transform of the data series and in particular compute the contribution corresponding to the signal period; we use both sine and cosine components because we have assumed that the phase of the signal is unknown. Then the sine and cosine components, $F^{(s)}$ and $F^{(c)}$, are given by

$$F^{(s)} = \frac{2}{QJ} \sum_{n=1}^{Q} \sum_{i=1}^{J} V_{(n-1)J+i} C_i^{(s)} \quad (2.4)$$

$$\text{and } F^{(c)} = \frac{2}{QJ} \sum_{n=1}^{Q} \sum_{i=1}^{J} V_{(n-1)J+i} C_i^{(c)}.$$

It is well known that for large values of Q (the number of operating periods available in the data series) the quantities $F^{(s)}$ and $F^{(c)}$ assume typical values in three particular cases:

a) If the data series represents a sinusoidal function with a period different from J then $F^{(s)}$ and $F^{(c)}$ tend to zero.

b) If the data series is characterised by a random distribution of values, namely if it represents a non-correlated non-periodic function, then $F^{(s)}$ and $F^{(c)}$ tend to zero.

c) If the data series represents a sinusoidal function with amplitude S and period J, but with a whatever phase φ, and if the sampling interval Δ is small, then $F^{(s)}$ and $F^{(c)}$ assume respectively the values

$$F^{(s)} = S\cos(\varphi) + \frac{2}{QJ} \sum_{n=1}^{Q} \sum_{i=1}^{J} \Phi_{(n-1)J+i} C_i^{(s)} \quad (2.5)$$

$$F^{(c)} = S\sin(\varphi) + \frac{2}{QJ} \sum_{n=1}^{Q} \sum_{i=1}^{J} \Phi_{(n-1)J+i} C_i^{(c)}.$$

In this case it is necessary to eliminate the effect of the second term of the right hand members of (2.5) and then the signal amplitude S can be computed as

$$S = (F^{(s)^2} + F^{(c)^2})^{1/2} \quad (2.6)$$

2.3. *Comparison between Fourier analysis and stacking procedure*

We can easily connect Fourier analysis and stacking procedure. In fact, rearranging the sums of (2.4) and using (2.1) we obtain

$$F^{(s)} = \frac{2}{J} \sum_{i=1}^{J} \left(C_i^{(s)} \frac{1}{Q} \sum_{n=1}^{Q} V_{(n-1=J+i} \right) = $$

$$= \frac{2}{J} \sum_{i=1}^{J} (C_i^{(s)} U_i) \quad (2.7)$$

and the similar expression for $F^{(c)}$.

Equation (2.7) clearly shows that the Fourier transform (2.4), computed with the QJ terms of the data series, is equal to the Fourier transform computed with the J values resulting from a stacking operation.

Namely Fourier analysis implies stacking, and some properties of the former are based on those of the latter. In particular the remarks following (2.4), both for a non-periodic and for a periodic series, derive from the analogous behaviour of the staking procedure (2.1).

Let us now suppose that a stacking procedure with operating period J is applied to a data series whose elements correspond to (1.2). In the Fourier transform the signal component is characterised by a monochromatic spectrum, while the noise component is represented by a continuous function of frequency. It is well-known that we have a finite amplitude of the Fourier transform of the signal for a given frequency, whereas the Fourier transform of the noise gives the relative amount of each sinusoidal component and we can define the amplitude density for a given frequency interval only.

As already pointed out, when the data series includes a sufficiently large number of operating periods, the results permit computation of the amplitude of the signal. In fact, the resulting amplitude of the non-periodic noisy component decreases when the number of available operating periods becomes greater. Unfortunately the stacking procedure is poorly efficient for the noise elimination in the period intervals around the value J and around its submultiples. On the other hand, the Fourier analysis appears to be useful since it can destroy more efficiently the high frequency noises and in any case it constitutes an additional operation with respect to the simple stacking, with a noticeably shortening of recording times.

The methods that we are going to explain in the next sections are based on the application of the Fourier analysis.

3. The «time dependent» Fourier analysis

The basic elements for the «time dependent» analysis are the Fourier components with period J computed for each operating period, *i.e.* for each row of (1.1)

$$A_n = \frac{2}{J}\sum_{i=1}^{J}(C_i^{(s)}V_{(n-1)J+i}) \text{ and}$$
$$B_n = \frac{2}{J}\sum_{i=1}^{J}(C_i^{(c)}V_{(n-1)J+i}), \quad n=1,\ldots,Q, \tag{3.1}$$

where $C_i^{(s)}$ and $C_i^{(c)}$ are the functions defined in (2.3).

A «time dependent» analysis is carried out computing the sum of A_n and B_n for the first m operating periods of the data series, *i.e.* computing the functions

$$F_m^{(s)} \equiv \frac{1}{m}\sum_{n=1}^{m} A_n \quad \text{and}$$
$$F_m^{(c)} \equiv \frac{1}{m}\sum_{n=1}^{m} B_n, \quad m=1,\ldots,Q. \tag{3.2}$$

Note that $F_Q^{(s)} = F^{(s)}$ and $F_Q^{(c)} = F^{(c)}$. We denote with $A(m)$ and $B(m)$ the absolute values of $F_m^{(s)}$ and $F_m^{(c)}$, namely

$$A(m) = \left|F_m^{(s)}\right| \text{ and } B(m) = \left|F_m^{(c)}\right|, \quad m=1,\ldots,Q. \tag{3.3}$$

In general A_n and B_n may assume casually positive or negative values, whereas $A(m)$ and $B(m)$ are positive by definition.

If the amplitude of the noise is constant during the data acquisition, the effects of the noise component can be correctly predicted by simple probabilistic laws and $A(m)$ and $B(m)$ have the following theoretical behaviour:

$$A(m) = \left|S_a \pm \alpha/\sqrt{m}\right| \text{ and } B(m) = \left|S_b \pm \beta/\sqrt{m}\right|, \tag{3.4}$$
$$m = 1,\ldots,Q,$$

where S_a and S_b represent, respectively, the sine and cosine components of the periodic signal, and the expressions $\alpha/\sqrt{m}$ and $\beta/\sqrt{m}$ describe the effect of noise on the sine and cosine components. Notice that we assume S_a, S_b, α and β to be positive quantities. Therefore the alternative sign in (3.4) corresponds to the possibility that signal and noise have concordant or opposite signs. We plot the theoretical diagrams of $A(m)/\alpha$ and $B(m)/\beta$ obtained from (3.4) in fig. 1.

These diagrams can be compared with the corresponding experimental diagrams and can be used to evaluate signal and noise amplitudes of the recorded data series. For each value of the ratios S_a/α and S_b/β listed in the caption to fig. 1,

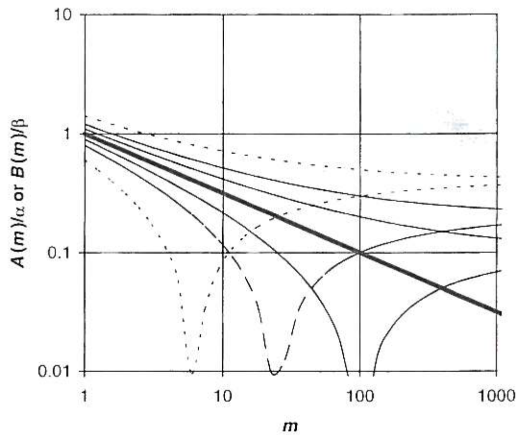

Fig. 1. Plot of $A(m)/\alpha$ and $B(m)/\beta$ as a function of the index m. The curves correspond to the following values of the signal-to-noise ratio (S_a/α or S_b/β): dotted line 0.4; dashed line 0.2; continuous thin line 0.1; continuous thick line 0.

we plot two diagrams, corresponding to the plus or minus sign in (3.4); in particular the upper diagrams are obtained for the plus sign, whereas the lower diagrams are obtained for the minus sign. These theoretical diagrams give the bounding curves for the diagrams obtained with real data for different SNR.

The theoretical model given by (3.4) crudely assumes that the amplitude remains approximately constant during data acquisition and shows that, in absence of signal, *i.e.* when pure noise is recorded ($S_a = S_b = 0$), $A(m)/\alpha$ and $B(m)/\beta$ are given by straight lines with a negative slope equal to $-1/2$.

The diagrams corresponding to the negative sign in (3.4) sharply decrease towards zero and then rise up to the limit value for great values of m. This happens because in this case the amplitude of the signal and the average amplitude of the noise have opposite signs; when their absolute values are approximately of the same order they tend to annihilate each other, so that $A(m)/\alpha$ and $B(m)/\beta$ tend to zero.

If we overlap a diagram obtained with real data over one of the master curves represented in fig. 1, we can estimate the SNR and, if the noise is approximately constant during the recording time, we can estimate the absolute values of the signal and the noise. We refer to this method of determination of the signal amplitude as «time dependent» analysis, since it is based on an examination of the functions $A(m)$ and $B(m)$, which are computed with the data of the first m operating periods of the whole recorded time series. These functions can be computed very easily on line, *i.e.* while data acquisition is going on.

Unfortunately, the condition of constancy of non-periodic noise during data recording poses a strong limitation on this analysis. The lack of this condition has the effect of changing the shape of the experimental diagrams, which are influenced not only by the amplitude and sign of the signal and the noise, but also by the time variations of the noise amplitude; this renders the comparison with theoretical master curves very difficult or even impossible. As a consequence, this method requires further improvement, with the aim of removing this difficulty. We suggest a solution in the next section.

4. The «time independent» Fourier analysis

The weak point of the «time dependent» analysis is the fact that $A(1)$ and $B(1)$ are computed with the first recorded operating period, and their values depend upon the SNR of this operating period, which in general is not representative of the average SNR of the whole recorded time series. This problem remains if we start with any of the operating periods, namely if we define $A(1)$ as $A(1) = |A_2|$, or $A(1) = |A_Q|$, or using any other single operating period. The same considerations apply to $A(2)$ and $B(2)$, since they are computed only by means of the first two periods, and so on for $A(m)$ and $B(m)$, with $m = 1, ..., Q$.

A simple choice to change this procedure and to reduce the dependence on the time variations of the noise level might be to compute $A(1)$ and $B(1)$ with several operating periods of the recorded series and obtain values that are more regular because they are influenced by the mean value of noise. The «time independent» analysis is based on the use of the whole data series, *i.e.* the computation of quantities similar to $A(m)$ and $B(m)$, $m = 1, ..., Q$, using all the available operating periods.

We call «super-averaged» functions the quantities $W_a(m)$ and $W_b(m)$, $m = 1, ..., Q$, defined as follows:

$$W_a(m) \equiv \frac{1}{Q}\sum_{n=1}^{Q}|H_a(m,n)| \equiv \equiv \frac{1}{Q}\sum_{n=1}^{Q}\left|\frac{1}{m}\sum_{s=0}^{m-1}A_{(s+n)\,\mathrm{mod}\,Q}\right|$$

$$W_b(m) \equiv \frac{1}{Q}\sum_{n=1}^{Q}|H_b(m,n)| \equiv \equiv \frac{1}{Q}\sum_{n=1}^{Q}\left|\frac{1}{m}\sum_{s=0}^{m-1}B_{(s+n)\,\mathrm{mod}\,Q}\right| \tag{4.1}$$

The functions $W_a(m)$ and $W_b(m)$ generalise the definition given by (3.3). In fact the quantities $H_a(m,n)$ and $H_b(m,n)$ represent the averages of Fourier transforms carried out on subsets of data composed of m operating periods; the number of these subsets is Q, strictly equal to the total number of recorded operating periods. Moreover, $H_a(m,1) = F_m^{(s)}$, $H_b(m,1) = F_m^{(c)}$ and

$$W_a(Q) = A(Q) = \left|F_Q^{(s)}\right| = \left|F^{(s)}\right| \quad \text{and}$$
$$W_b(Q) = B(Q) = \left|F_Q^{(c)}\right| = \left|F^{(c)}\right|. \tag{4.2}$$

In other words, the last value of the «super-averaged» functions represents the absolute value of the average of the components for all the operating periods.

The computation of the quantities $H_a(m,n)$ and $H_b(m,n)$ is based on the Fourier analysis for the operating period $(s+n)$ mod Q, which means that if $s+n > Q$, then $A_{(s+n)\mathrm{mod}Q} = A_{s+n-Q}$ and $B_{(s+n)\mathrm{mod}Q} = B_{s+n-Q}$ in (4.1). This is equivalent to stating that the results of the Fourier analysis for each operating period lay on a circle and the sums that define the quantities $H_a(m,n)$ and $H_b(m,n)$ are performed along the circle itself.

The following formula shows how this rule work for a particular example, when $Q = 32$:

$$H_a(8,27) = \tag{4.3}$$
$$= (A_{27}+A_{28}+A_{29}+A_{30}+A_{31}+A_{32}+A_1+A_2)/8.$$

The proposed «time independent» analysis consists of the following steps:

a) Computation of the sine and cosine Fourier transforms A_n and B_n for each operating period, *i.e.* for the data in each row of (1.1), using (3.1). It is important to point out that these quantities have positive or negative signs in a random way if the noise prevails over the signal, whereas their sign is the same as that of the signal, if the latter prevails over the noise.

b) Computation of the «super-averaged» functions $W_a(m)$ and $W_b(m)$, $m = 1,...,Q$. Note that it may be sufficient to compute $W_a(m)$ and $W_b(m)$ for a limited number of values of the independent variable m to obtain good results. In fact, since the two functions are usually smooth and plotted with log-log scales, the calculation of the whole function of m should not be necessary.

We stress that according to (4.1), the index n denotes the first operating period used for the computation of $H_a(m,n)$ and $H_b(m,n)$, whereas the remaining $(m-1)$ elements of the sum are those following in time order. However the time order may be disregarded, as it is shown, for instance by (4.3). In particular we could define the quantities $H_a(m,n)$ and $H_b(m,n)$, introducing all the possible permutations of the operating periods. We avoid this further generalisation, because it does not add any sensible improvement to the «time independent» analysis.

From (4.1) it appears that the computation of $H_a(m,n)$ and $H_b(m,n)$ consists of the algebraic summation of the Fourier analysis of the operating periods, similarly to the «time dependent» analysis, but only the absolute values of these sums are used for the successive operations and the determination of the «super-averaged» functions. A difference between the «time dependent» procedure and the «time independent» one is that, for each value of m, more groups of operating periods are used for the «time independent» analysis than for the «time dependent» analysis; in particular if we consider operating periods in time order, for each value of m, we use the maximum number of groups of operating periods for computing the «super-averaged» functions, namely mQ.

c) The «super-averaged» functions $W_a(m)$ and $W_b(m)$ define two diagrams concerning the sine and cosine components of the recorded data set which permit evaluation of the signal amplitude according to the criteria described in the following section.

5. Determination of signal amplitude with «super-averaged» functions

We distinguish three cases. The first case concerns an absence of periodic signal. According to probability principles, the noise, if it is perfectly random and even if its amplitude is not constant with recording time, gives diagrams of «super-averaged» functions, namely $W_a(m)$ and $W_b(m)$ as functions of m, which are straight lines with a negative slope equal to $-1/2$ in log-log scales.

The second case concerns the presence of a periodical signal without any noise. In this case the diagrams are straight lines parallel to the abscissa axis. The two constant values of «super-averaged» functions correspond respectively to the sine and cosine components of the periodical signal.

The third and most interesting case regards data series for which the mean oscillation amplitude of the noise varies with recording time and the SNR is much smaller than 1. In this case the signal is completely masked and it is hardly detectable with the stacking procedure or the «time dependent» analysis, if the number of operating periods is not very high. In this case the variation of noise amplitude with time might be important. Therefore the last contribution of the right hand side of (1.2) has to be removed; this can be done in an effective way assuming that this quantity is represented by a linear trend between the measured values $V_{(n-1)J+1}$ and V_{nJ+1}. Our experience shows that removing this linear trend from each operating period is sufficient to eliminate the effects of $\Phi_{(n-1)J+i}$ on the «super-averaged» functions.

Both diagrams derived by the functions (3.4) present typical shapes. For $m = 1$, $W_a(1)$ and $W_b(1)$ are influenced mostly by the noise but also by the signal amplitude and by the phase of the signal. Because of the noise prevalence and according to the already said probability principles, for $m > 1$ the values of the two «super-averaged» functions decrease so that the corresponding diagrams, in log-log scale, generally show a negative slope, whose value is smaller or greater than $-1/2$, which is the signature of lack of signal.

The difference in the slope of the «super-averaged» functions (greater or smaller than $-1/2$) in presence of a signal is caused by the signs of the average value of the noise components and the signal. In fact, if we recall that (3.1) and (3.2) involve only linear operators and take into account (1.2) and (4.3), we can easily prove that $W_a(Q)$ and $W_b(Q)$ correspond to the absolute values of the algebraic sum of, respectively, the sine and cosine components of the signal and of the average noise. Therefore when the average of a noise component evaluated on the basis of the whole data series is opposed to the value of the same component of the signal, then the «super-averaged» function corresponding to that component tends to zero for increasing m, so that its diagram in log-log scale shows a slope smaller than $-1/2$. On the other hand, if a signal component has the same sign as the corresponding component of the average noise and if a sufficiently long record has been acquired, *i.e.* if we have recorded many operating periods, then we observe a characteristic behaviour of the «super-averaged» functions. In particular there is a definite value of m, say m', such that $W_a(m) = W_a(m') = W_a(Q)$, for every $m = m'$, ..., Q, or an analogous property for the cosine component. In other words the «super-averaged» function becomes exactly constant. This is a very simple and useful criterion to establish the presence of a signal in the acquired data set. It is easy to check from (4.1) that $W_a(m) = W_a(m')$, $m = m', ..., Q$, if all the quantities $H_a(m', n)$, $n = 1, ..., Q$, have the same sign.

Finally the determination of the signal components can be obtained by subtracting tentative arbitrary values of the signal components from A_n and B_n; when the tentative values correspond to the correct values of the signal components, the modified Fourier transforms for each operating period would include only the effect of the random noise and the diagrams of «super-averaged» functions should approximate a straight line with angular coefficient $-1/2$ in log-log scale. Therefore the determination of the signal components is based on the fitting of the «super-averaged» functions with such a simple master curve.

6. Examples with synthetic and real data

We show some results with synthetic data in fig. 2a-c. In particular we have generated synthetic data series for $Q = 32$ as follows: A_n, $n = 1, ..., Q$, have been computed by adding pseudo-random uniformly distributed numbers which simulate noise to an assigned signal value, S. We have modelled several different realisations of the synthetic random noise and the results are qualitatively the same as those represented in fig. 2a-c which correspond to a noise with mean value equal to 0.1 and standard devi-

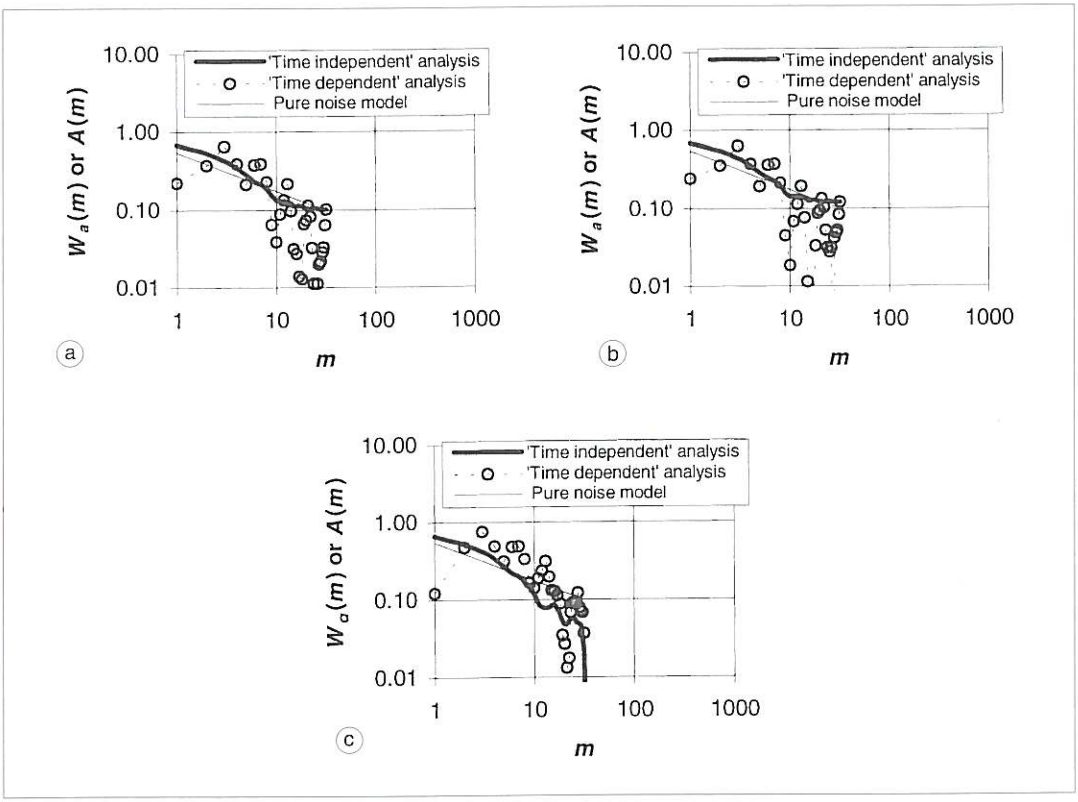

Fig. 2a-c. Plots of $W_a(m)$ and $A(m)$ for synthetic data sets. a) Pure noisy data; b) signal superimposed on the noise; c) signal opposed to the noise.

ation equal to 0.8. We consider three cases, corresponding to different values of the signal: $S = 0$, $S = 0.02$ and $S = -0.1$. For each case in fig. 2a-c we plot both the «super-averaged» function, $W_a(m)$, and the function $A(m)$; for the reader's convenience we plot in the same figure a straight line which corresponds to the theoretical model given by (3.4).

Figure 2a shows the results for the case of pure noise, *i.e.* when $S = 0$; it is easily seen that the diagram of the «super-averaged» function approximates a straight line much better than the function $A(m)$, which shows an erratic trend.

Figure 2b shows the results when $S = 0.02$, *i.e.* when the signal is 1/5 of the average value of the noise and the ratio between the signal and the standard deviation of noise is 1/40. In this case the «super-averaged» function becomes constant for $m \geq m' = 27$. In this case the «time dependent» analysis does not allow us to obtain a confident estimate of the signal, nor it can indicate clearly if the signal can be revealed by the recorded data set. In fact the behaviour of the function $A(m)$ is still quite erratic so that it is difficult to infer whether signal is present and to evaluate its amplitude.

Figure 2c shows the results for $S = -0.1$, *i.e.* when the signal and the average noise are opposed. We see that the «super-averaged» function decreases rapidly, namely with a slope smaller than $-1/2$, *i.e.* it decreases by a factor of ten in less than two decades for m.

These results show the advantage of the «time independent» analysis with respect to «time dependent» analysis. In particular we see that the analysis of the «super-averaged» functions allows us to obtain several information. Moreover fig. 2a-c shows that the «super-averaged» functions are very smooth, so that in the field it is not necessary to compute the whole function for every value of m, but it suffices to compute a limited number of values. However the «super-averaged» functions presented in this paper have always been computed for all the values of m, for the sake of completeness.

We now show the results corresponding to the application of the Fourier analysis to the real data set represented in fig. 3 which correspond to $\Delta = 1$ s, $J = 20$ and $Q = 40$. In particular we plot both the gross data, recorded in the field (thick line in fig. 3) and the data after the reduction for removing the long-period trend that corresponds to the term $\Phi_{(n-1)J+i}$ of (1.2). The stacking procedure for this example seems to give good results if we look at the plot of fig. 4; in fact the stacking procedure shows a coherent trend with a signal phase φ close to 180°. On the other hand, the plot of the error bars shows that the determination of the signal amplitude from the stacking procedure is quite uncertain.

The results of the Fourier analysis are shown in fig. 5a,b. Since the phase of the signal is close to 180°, the sine component should be less than the cosine component. This is clearly shown by fig. 5a, which refers to the sine component; the results of the «time independent» Fourier analy-

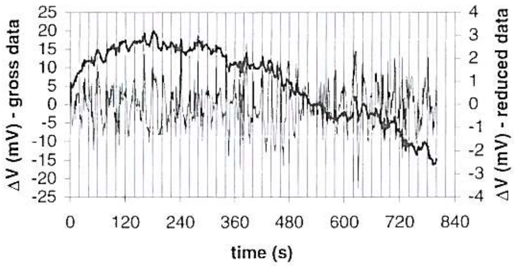

Fig. 3. Real time series of potential differences. Thick line = gross data (values on the left axis). Thin line = data after removing the long-period trend (values on the right axis). Vertical bars separate the operating periods.

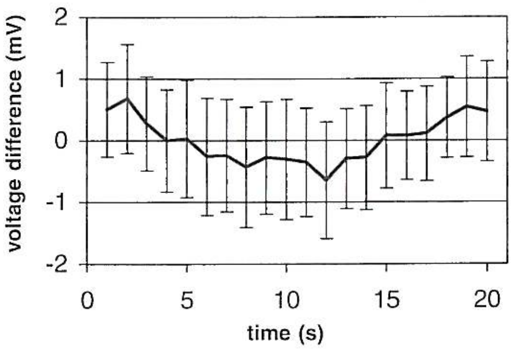

Fig. 4. Results of the stacking procedure for the real data set of fig. 3. The length of the error bars corresponds to twice the standard deviation.

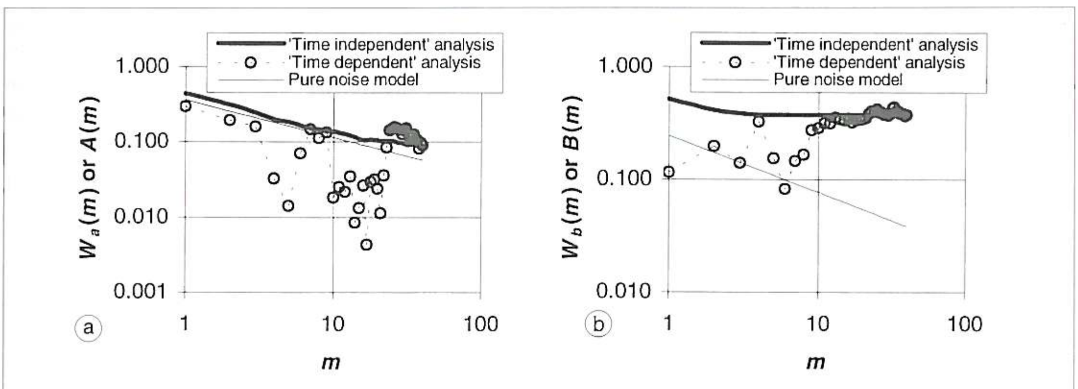

Fig. 5a,b. Plots of the results of the Fourier analysis for the real data set of fig. 3. a) Sine component; b) cosine component.

sis show that this component is mainly affected by noise, but nevertheless permits us to evaluate its amplitude. In particular $W_a(m)$ becomes constant for $m \geq m' = 38$, whereas the results of the «time dependent» analysis are less clear and more ambiguous. On the other hand the «super-averaged» function $W_b(m)$ for the cosine component becomes constant for $m \geq m' = 7$; this gives clear evidence that the data series contains a good signal.

The importance of the use of the «super-averaged» function for this example is twofold: first, the «super-averaged» function permits us to evaluate the signal amplitude; second, it confirms the presence of a periodic signal in the data series and increases the reliability of the results of the stacking procedure.

7. Conclusions

We have shown that several techniques for extracting electrical signal from noisy data series are based on refinements and developments of the stacking procedure. In particular, we have analysed this for the classical technique of Fourier analysis.

The «time dependent» Fourier analysis cannot be efficient if a large number of operating periods has not been recorded and if the noise level varies during data recording. Therefore we have introduced the «time independent» Fourier analysis, which is based on the use of the classical Fourier analysis and of the «super-averaged» functions. With this technique we do not evaluate the signal components with the classical technique of periodogram estimation but we introduce functions which use the results of Fourier analysis for all the available operating periods in the data series. This permits us to obtain estimates of the electrical signal in a more reliable way than with other techniques and, in particular, also for unfavourable signal-to-noise ratios.

Acknowledgements

This research was financially supported by LEMI Project (MURST ex-40% funds) and by CNR (Contribution 96.00293.CT05). We thank the reviewers for their helpful comments.

REFERENCES

ALFANO, L., E. CARRARA, G. PASCALE, A. RAPOLLA and N. ROBERTI (1982): Analysis procedure and equipment for deep geoelectrical soundings in noisy areas, *Geothermics*, **11**, 269-280.

CIMINALE, M. and D. PATELLA (1982): On the determination of the maximum likelihood power of very noisy voltage records in geoelectrical exploration, *Boll. Geofis. Pura Appl.*, **24**, 205-210.

LAPENNA, V., C. SATRIANO and D. PATELLA (1987): On the methods of evaluation of apparent resistivity under conditions of low message-to-noise ratio, *Geothermics*, **16**, 487-504.

LAPENNA, V., M. MACCHIATO, D. PATELLA, C. SATRIANO, C. SERIO and V. TRAMUTOLI (1994): Statistical analysis of non-stationary voltage recordings in geoelectrical prospecting, *Geophys. Prospect.*, **42**, 917-952.

LODDO, M. and D. PATELLA (1977): Spectral analysis of voltage measurements for the interpretation of dipole geoelectrical soundings in low resistivity geothermal areas, *Geothermics*, **6**, 227-235.

OPPENHEIM, A.V. and R.W. SCHAFER (1989): *Discrete-Time Signal Processing* (Prentice Hall, Englewood Cliffs, NJ), pp. 879.

TELFORD, W.M., L.P. GELDART and R.E. SHERIFF (1990): *Applied Geophysics* (Cambridge University Press, Cambridge), 2nd edition, pp. 770.