A Magnetotelluric Investigation of Geoelectrical Dimensionality and Study of the Central Betic Crustal Structure
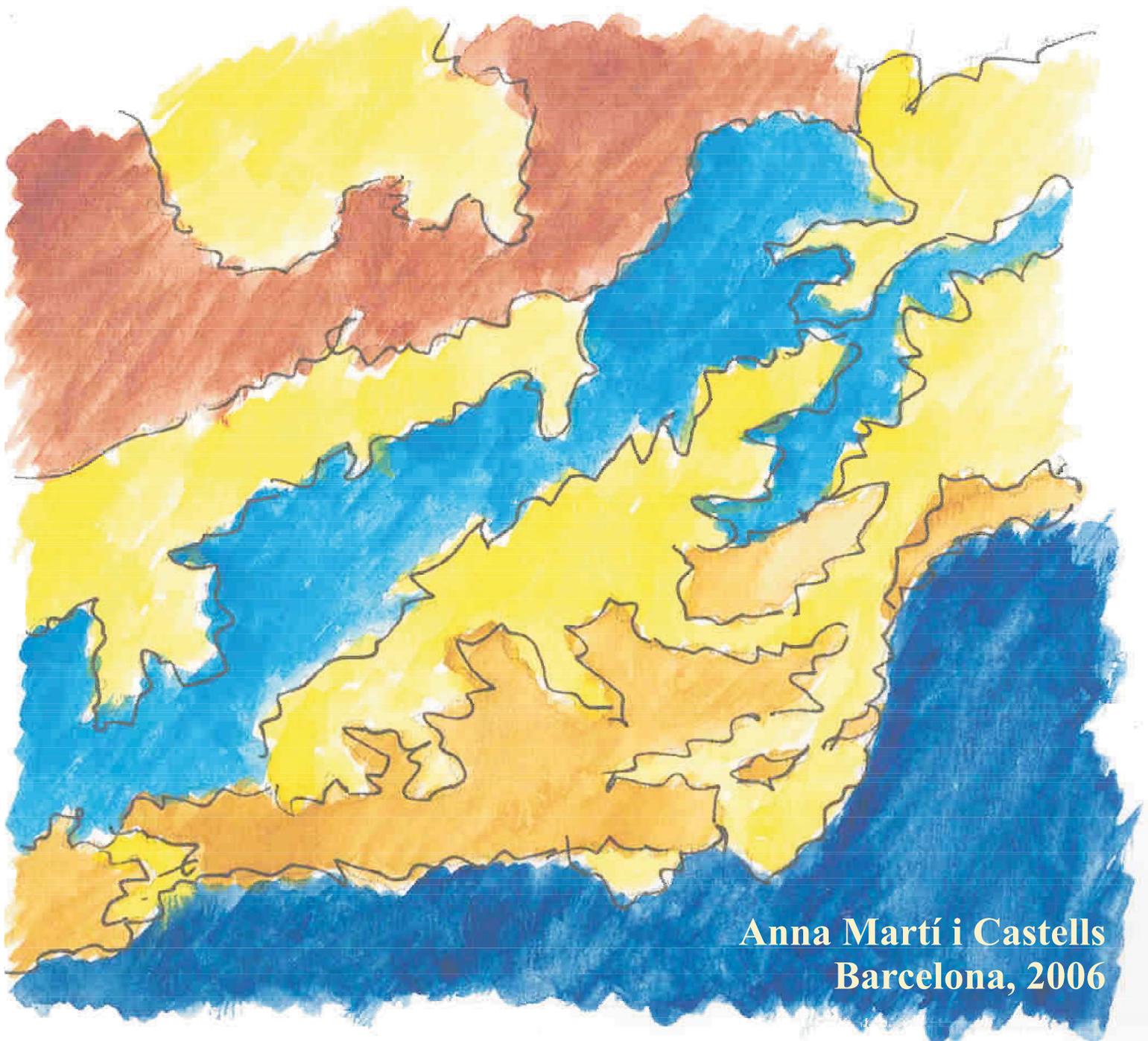
**Anna Martí i Castells
Barcelona, 2006**

Part II. Methodological Contributions to Geoelectric Dimensionality

Chapter 3: Error Analysis and Threshold Values in WAL Rotational Invariants

This chapter develops new procedures which were carried out on the error estimation of the invariants and the choice of their threshold values. To this end, it proposes a new method that allows the characterisation of geoelectric dimensionality in real situations.

Two examples, with different levels of errors, are used to illustrate the problems that one must overcome in order to obtain an accurate dimensionality analysis, and thus, further modelling and interpretation.

With the aim of introducing a tool to automatically perform dimensionality analysis from a dataset, a program was developed based on WAL criteria, considering noise in the data and its implications for dimensionality determination.

Part of this work can also be found in Martí *et al.* (2004).

3.1. Dimensionality Criteria Using Real Data

The main problem when WAL invariants criteria are implemented on real, therefore noisy data is that the geoelectric dimensionality may be found to be 3D. Although, other evidence suggests a 1D or 2D interpretation would be valid for modelling. This is because invariant values for real data are, in general, never precisely zero. Weaver *et al.* (2000) address this problem by introducing a threshold value, beneath which the invariants are taken to be zero. The threshold value they suggest is 0.1, which, although subjective, has been tested using a synthetic model with 2% noise. Since experimental data usually have a higher percentage of

error (up 30% in some cases) which propagates to the invariants, it is necessary to redefine this threshold value, taking into consideration the invariant values and their errors. To date, no work has been done on the estimation of WAL invariant errors.

Using WAL criteria with the threshold defined, if the dimensionality obtained is 2D or 3D/2D (cases 2, 3a and 4 in table 2.2), the strike directions and distortion parameters must also be estimated with their errors.

To address these matters, different tests were performed to estimate the invariants, related parameters, and their errors, and to choose an optimum threshold value.

3.2. Estimation of the Invariants and their Errors

The values and errors of the invariants and related parameters are dependant upon the values of the MT tensor components and their errors, which result from the estimation of the transfer functions in the spectral domain, after time series processing. In standard processing methods, these errors are obtained as variances, var(M_{ij}), after assuming that noise is independent of the signals and stationary, and, consequently, that the components of $\underline{M}$ are statistically independent (Chapter 1, section 1.7, eq. 1.41). The errors of the real and imaginary parts of the MT tensor components have the same value and are determined as the square root of its variance:

$$\delta(\operatorname{Re} M_{ij}) = \delta(\operatorname{Im} M_{ij}) = \delta(M_{ij}) = (\operatorname{var}(M_{ij}))^{1/2}. \tag{3.1}$$

Invariant errors reflect dimensionality uncertainties due to data errors. Hence, it is important to properly estimate both the invariants and their errors to have a consistent picture of how well determined the dimensionality of a MT tensors is. Similarly, the errors of the related parameters (strike angles and distortion parameters) indicate incorrect determinations of certain types of dimensionality.

Because of the non-linear dependence of the invariants and related parameters on the MT tensor components, the relationships between their errors are highly complex. Given the role that the errors play in the determination of the dimensionality, three approaches were tested in order to obtain a high level of confidence in the estimation of the invariants and related parameters and errors: **a**. Classical error propagation, **b**. Random Gaussian noise generation and **c**. Resampling methods. The results and stability using these approaches were compared to determine which is the most suitable for the particular parameters.

3.2.1. Error Estimation Approaches

a. Classical Error Propagation

The values of the invariants are computed directly from expressions 2.10 to 2.16, and are hereafter referred to as true values (I).

For small errors, the uncertainty of any function $y = f(x_1, x_2,..., x_n)$ can be obtained from a Taylor expansion in terms of the errors ($\delta x_1, \delta x_2,...\delta x_n$) of the estimated variables $x_1, x_2,..., x_n$. Using a first order expansion, the error (δI_k) of each invariant (I_k) and the errors of the strike and distortion parameters can be expressed as functions of partial derivatives (of the corresponding invariant expressions) and the errors of the real and imaginary parts of the components of $\underline{M}$:

$$\delta I_k = \left(\sum_{i=1}^{2} \sum_{j=1}^{2} \left(\left(\left(\frac{\partial I_k}{\partial (\mathrm{Re}\,M_{ij})} \right)^2 + \left(\frac{\partial I_k}{\partial (\mathrm{Im}\,M_{ij})} \right)^2 \right) \cdot (\delta(M_{ij}))^2 \right) \right)^{1/2}. \tag{3.2}$$

These expressions were obtained as functions of ξ_i and η_i, and are summarised in Appendix A.

b. Random Gaussian Noise Generation

This statistical approach is an alternative to the first, which can fail when the errors in $\underline{M}$ are large. The problem lies in the lack of knowledge about the statistical distribution of $\underline{M}$, since the only available data are the means and variances of their components.

Some authors assume a Gaussian distribution of $\underline{M}$ in order to compute new parameters and study their stability (Jones and Groom, 1993; Weaver *et al.*, 2000), or to obtain their probability functions and confidence limits (Lezaeta, 2002), both from synthetic and real data.

Thus, one way to estimate the invariants and their errors is to generate a set of n possible values of $\underline{M}$ components, M_{ij}^l $(l = 1, 2,..., n)$, assuming Gaussian noise around their true values, with the variances of the MT tensor components:

$$\mathrm{Re}\,M_{ij}^l = \mathrm{Re}\,M_{ij} + \mathrm{rndG}(\delta M_{ij}); \quad l = 1, 2,..., n, \tag{3.3}$$

$$\mathrm{Im}\,M_{ij}^l = \mathrm{Im}\,M_{ij} + \mathrm{rndG}(\delta M_{ij}); \quad l = 1, 2,..., n, \tag{3.4}$$

where $\mathrm{rndG}(\delta M_{ij})$ are pseudo-random numbers that follow a Gaussian distribution with an average of zero and a standard deviation of δM_{ij}.

From this set of values, the corresponding n possible values of the invariants can be obtained:

$$I_k^l = f(\operatorname{Re} M_{ij}^l, \operatorname{Im} M_{ij}^l; i = 1, 2; j = 1, 2); \quad l = 1, 2, ..., n \qquad . \tag{3.5}$$

After obtaining this set of n realizations for each invariant at a determined site and period, its mean value (I_k') and standard deviation ($\sigma_{I_k'}$), which is taken as the error, are estimated as:

$$I_k' = \frac{1}{n} \sum_{l=1}^{n} I_k^l , \tag{3.6}$$

$$\sigma_{I_k'} = \sqrt{\frac{1}{n} \sum_{l=1}^{n} \left(I_k' - I_k^l \right)^2} . \tag{3.7}$$

The value of n must be chosen to avoid biases between the true value (I_k) and the mean value (I_k') of the invariant. One criterion to fix the optimum value of n is to ensure that the bias between I_k and I_k' is not greater than the standard deviation (i.e. not a significant bias):

$$\left| I_k - I_k' \right| \le \sigma_{I_k'}. \qquad \text{(bias condition)}. \tag{3.8}$$

The same procedure can be used to obtain the mean value and standard deviation of the strike direction and distortion parameters.

c. Resampling Methods

Resampling methods are statistical procedures used to obtain confidence intervals for parameters whose probability distributions are unknown (Efron and Tibshirani, 1998). They are based on the construction of hypothetical populations derived from the measured data. No statistical distribution is assumed, although this distribution is preserved. It follows that the new populations can be analyzed individually to see how the statistics depend on random variations of the data.

Two well-known resampling methods are the jackknife (Lupton, 1993), and the bootstrap (Efron, 1979). These two methods differ in how the hypothetical populations are constructed for an original dataset of n datapoints. The Jackknife method constructs n populations of n-1 data points, omitting one different point in each. Some robust data processing in MT are based on this technique (Jones and Jödicke, 1984; Thomson and Chave, 1991). In the

bootstrap method, also known as 'resampling with replacement', a large number of populations (or resamples) with n elements each can be constructed, allowing duplicates and triplicates of some points while allowing the absence of others.

The bootstrap method searches for a good approximation of the unknown population distribution, such that the estimated parameters also resemble those of the original population. The bootstrap method has been demonstrated to provide optimal estimates of the distribution of many statistics under a wide range of circumstances (Efron and Tibshirani, 1998). For this reason, the bootstrap approach was chosen to test the estimation of the invariants and their errors.

The most common application of the bootstrap method is the nonparametric-bootstrap, in which all data points are given the same weight. In this case, the number of resamples (N) necessary to approximate the theoretical sampling distribution is on the order of $n \cdot (\log n)^2$ (Babu and Singh, 1983). For every bootstrap sample, a statistical estimator and its confidence intervals are computed, recreating an empirical sampling distribution of the estimator.

The estimation of the invariants and related parameters and their errors using the bootstrap method can be done departing from a set of n elements as the initial population. These elements are generated as random Gaussian noise (as in approach **b**) (one set for each invariant or parameter). N resamples with n elements each are then constructed. The mean and standard deviation of the parameter is obtained from the mean and standard deviation of each of the bootstrap populations.

This application of the bootstrap method has the limitation that the original datasets are constructed assuming a probability distribution (Gaussian, from approach **b**), and are not raw data. A more rigorous use of bootstrap could be done if the original cross-spectra of the MT tensor were available, such that n estimates of the invariants could be obtained, and thus be used as the initial population. In many cases, it is not possible to do so, since the only available data after time series processing are the final estimates of the MT transfer functions, their coherences and related parameters. However, it is worth testing as an additional alternative to estimate the invariant values and their errors.

The following subsections present the comparisons and results of the estimations of the invariants and their errors using the three approaches described. Two data examples are used, for which invariants I_3 to I_7 and Q (the ones used in the dimensionality determination) were considered.

3.2.2 Example 1

The first example is site 85_314 from the COPROD2 dataset (Jones, 1993) acquired in Saskatchewan and Manitoba (Canada). It has been widely employed by the MT community to test 2D inversion codes, since the data clearly display 2D features, with a well determined NS strike direction (Jones *et al.*, 2005). The location and MT responses at this site are presented in Appendix B. The measured MT tensor components, with periods ranging from 0.002 s to 1000 s, are of good quality. Diagonal components are one or two orders of magnitude smaller than the non-diagonal ones, this ratio decreasing as the period increases. The average level of noise is 5%, with a minimum of 1% between 5 s and 500 s. This minimum is a result of the high quality of the acquisition system, Phoenix MT-16, within this period band.

Determination of n and N values in Random Gaussian Noise and Bootstrap approaches

Prior to the estimation of the parameters and their errors from approaches **b** (Random Gaussian Noise Generation) and **c** (Bootstrap), some tests were carried out on the number of realizations (n) (approach **b**) and on the size of the samples and the number of resamplings (approach **c**), using this data example.

- Random Gaussian Noise

Several values of n (n=10, n=100 and n=1000) were tested to evaluate the dispersion of the invariants, their means and standard deviations, and were compared to their true values and classical errors in order to check whether the bias condition is satisfied or not. Invariants I_5 and Q were chosen as representatives of simple and complex dependencies on $\underline{M}$ components, respectively.

Figure 3.1 illustrates the realizations of I_5 and Q, their means and their standard deviations for three periods, T=0.67 s, T=0.0208 s and T=21.32 s, with high, medium and low noise levels respectively in the MT tensor components, using n=10 and n=100. These are compared to the true values and their errors obtained using approach **a** (n=1000 is not shown for simplicity and because it presents similar results to those of n=100). The numerical values of these estimations, including n=1000 and the results from the bootstrap, are shown in table 3.1. For these three periods, the following can be inferred:

At T=0.67 s (high noise level: 10% error in $\underline{M}$ non-diagonal components and up to 200% in the diagonal ones), I_5 and Q present a large dispersion and consequently large error bars. Errors of I_5 are similar to the classical errors and those of Q are smaller, independent of the value of n. The realisations and statistical estimations of Q are up-biased and, for small values of n (n=10), this bias is, in fact, significant (equation 3.9 is not satisfied). The estimations of I_5 are biased also, although, due to the small values, it is only appreciable numerically (Table 3.1).
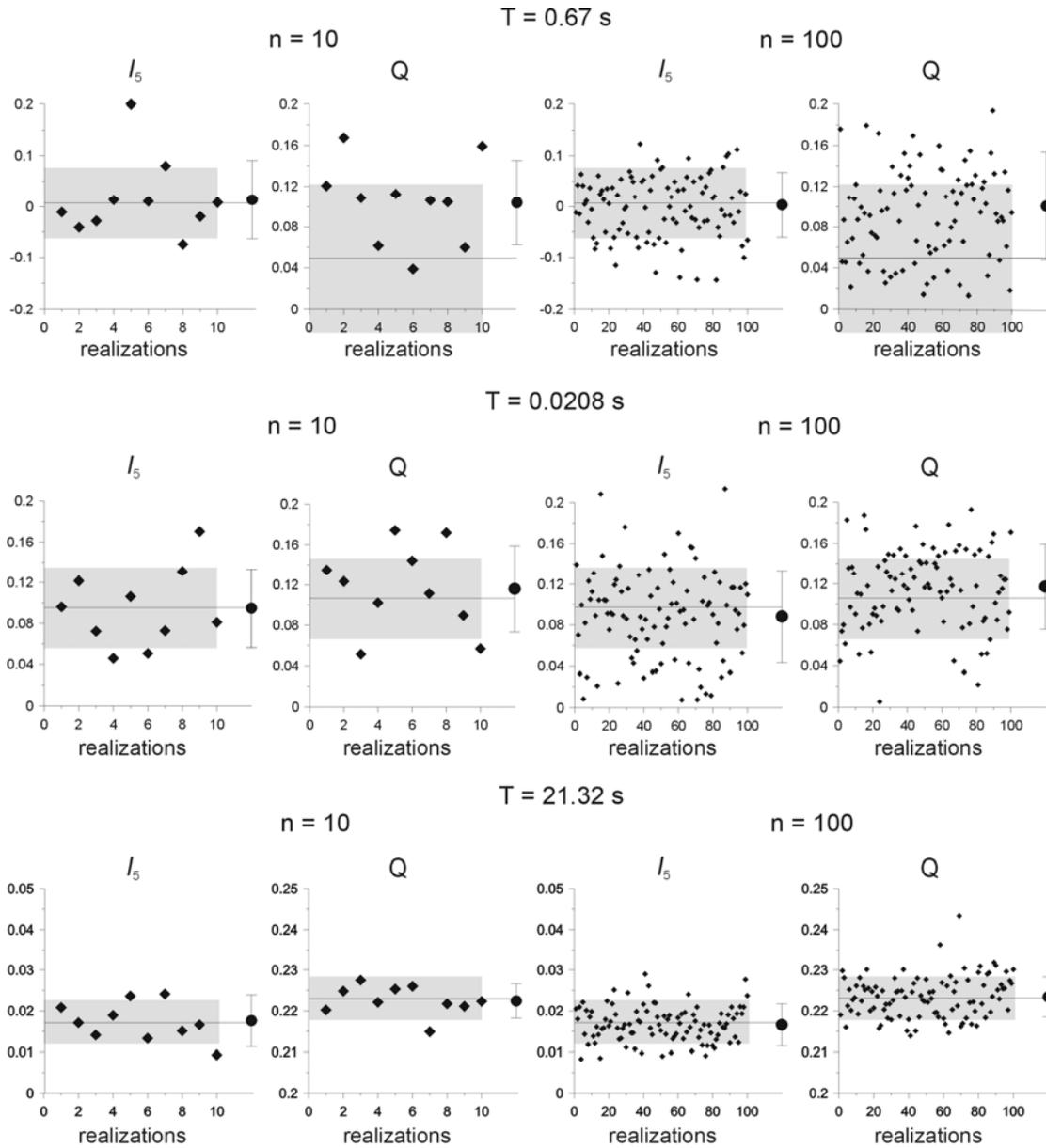
Figure 3.1: Plots of n=10 and n=100 realizations (diamonds) of invariants I_5 and Q at periods T=0.67 s (high noise level, upper panel), T=0.0208 s (medium noise level, middle panel) and T=21.32 s (low noise level, lower panel) from site 85_314 (COPROD2 dataset). The mean and standard deviations of these realizations are displayed at the right end of each plot (dot with error bar). Horizontal lines at each plot indicate the true value of the invariant. The shaded areas correspond to the confidence intervals, as obtained from classical error propagation (approach **a**).

At T=0.0208 s (medium noise level: 5% error in $\underline{M}$ non-diagonal components and up to 100% in the diagonal ones), the noise level of the invariants is about 40%. Both I_5 and Q show a uniform distribution of the random values around the true values, with similar errors, and there is not an appreciable bias between the true and statistical values.

At T=21.32 s (low noise: 0.5% error in the non-diagonal components, and 8% in the diagonal ones) the dispersion of both I_5 and Q is much smaller (note the different vertical scales on the plots). Error bars are similarly small and in all cases the bias condition (equation 3.9) is satisfied.

		T=0.67 s		T=0.0208 s		T=21.32 s	
		I_5	Q	I_5	Q	I_5	Q
True		0.0067 ± 0.0696	0.0492 ± 0.0717	0.0972 ± 0.0422	0.1057 ± 0.0428	0.0171 ± 0.0052	0.2231 ± 0.0052
Random	n=10	0.0140 ± 0.0769	0.1039 ± 0.0414 *	0.0935 ± 0.0446	0.1148 ± 0.0432	0.0179 ± 0.0054	0.2211 ± 0.0051
	n=100	0.0034 ± 0.0639	0.1005 ± 0.0529	0.0913 ± 0.0407	0.1229 ± 0.0396	0.0170 ± 0.0048	0.2230 ± 0.0045
	n=1000	0.0034 ± 0.0639	0.1005 ± 0.0529	0.0966 ± 0.0414	0.1145 ± 0.0399	0.0172 ± 0.0049	0.2231 ± 0.0050
Bootstrap	n=10 N=10	0.0070 ± 0.0668	0.1098 ± 0.0392 *	0.0979 ± 0.0599	0.0824 ± 0.0337	0.0178 ± 0.0050	0.2203 ± 0.0050
	n=100 N=400	0.0036 ± 0.0637	0.1009 ± 0.0527	0.0996 ± 0.0397	0.1094 ± 0.0394	0.0171 ± 0.0047	0.2230 ± 0.0044
	n=1000 N=9000	0.0036 ± 0.0637	0.1009 ± 0.0527	0.0967 ± 0.0443	0.1150 ± 0.0405	0.0170 ± 0.0052	0.2231 ± 0.0048

Table 3.1: Estimations of I_5 and Q and their errors at two periods from the COPROD2 dataset. True: True values and classical error propagation. Random: means and standard deviations from Random Gaussian generation (approach **b**), for n=10, n=100 and n=1000. Bootstrap: N resamplings of the initial samples from the Random approach. *****: Cases with a significant bias between statistical and true values.

At all periods, the values of $\sigma_{I_k'}$ are similar for n=100 and n=1000. Although errors are not much larger, n=10 is not representative enough of all possible values of an invariant. Tests with other invariants and periods, using a broader range of values of n show a stable pattern of $\sigma_{I_k'}$ using n=100 to n=1000. Consequently, values of n=100 or greater are recommended to perform approach **b**.

- Bootstrap

A similar test was carried out for the same three periods with different values of n (size of the initial samples), and N (number of resamples $\approx$ n·(log n)^2). The samples of I_5 and Q generated in the previous test, with n=10, n=100 and n=1000, were resampled N≈n·(log n)^2 times (N=10, N=400 and N=9000 respectively) and their means and standard deviations estimated.

The results (both statistical values and standard deviations), shown in Table 3.1, do not differ significantly to the ones with the corresponding value of n from approach **b**, especially when the noise level is moderate or low (T= 0.0208 s and T=21.32 s).

At T=0.67 s, with n=10, there is also a significant bias between the bootstrap and the true values of Q. This bias disappears for n=100 and n=1000.

In this approach, n=100 and n=1000 (and, consequently $N \approx 400$ and $N \approx 9000$) are valid to estimate the invariants and their errors using the bootstrap approach, although $N \approx 400$ is sufficient and is the value that will be used in the following comparison.

Comparison of the three approaches

Figure 3.2 shows the invariant values and error bars after using the three approaches **a**, **b** and **c**, at all the recorded periods of site 85_314. Approach **b** uses n=1000 and approach **c** utilises N=400, from an initial sample containing n=100 elements each.

Using the three approaches, invariants I_3, I_4, I_5 and I_6 present similar values and errors. Error bars are proportional to the noise level of the recorded data (note small error bars between T=5 s and T=500 s), and no appreciable biases are observed between the true values, **a**, and the statistical values of approaches **b** and **c**.

I_7 is characterised by large error bars, which only decrease between 5 s and 500 s, when Q becomes large, that is, when I_7 (which is inversely proportional to Q) is well determined. The differences between the estimations of I_7 using the three approaches are evident, especially seen in the sizes of the error bars. From approach **a** the error bars present an irregular distribution, with extremely large values at certain periods. Approaches **b** and **c** provide a smoother distribution of the errors, but introduce important biases with respect to the true values (note the differences in I_7 between (g), (h) and (i) in Figure 3.2). Invariant Q, on the contrary, presents similar estimations using any of the three methods, with small values and error bars, proportional to the data noise level.
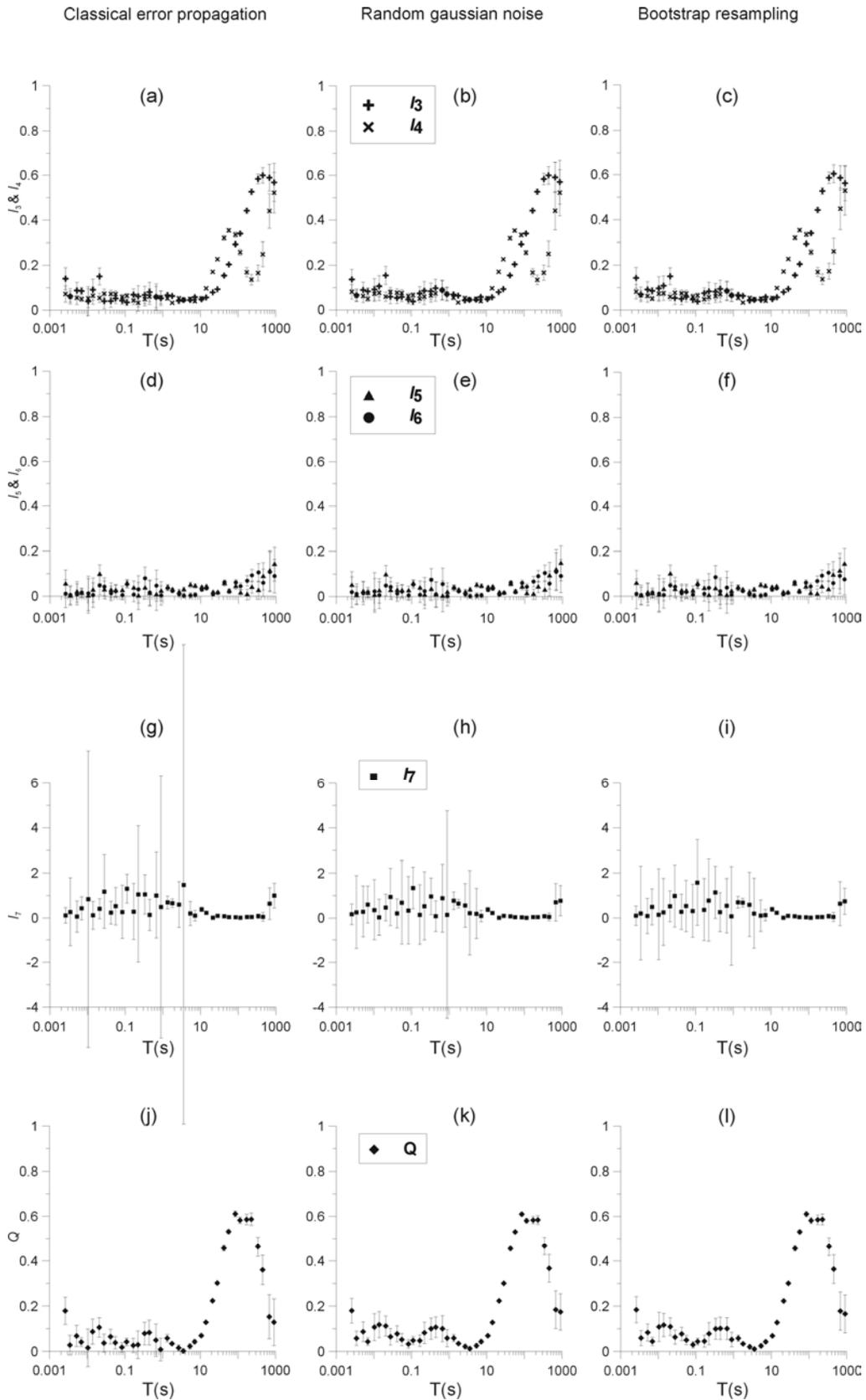
Figure 3.2: Invariant values at site 85_314 from the COPROD2 dataset and their errors computed in three ways: classical error propagation, Random Gaussian Noise (1000 realizations) and bootstrap resampling (400 realizations from an original sample with 100 elements for each invariant). I_3 and I_4 (a), (b) and (c); I_5 and I_6 (d), (e) and (f), I_7 (g), (h) and (i) and Q (j), (k) and (l).

3.2.3 Example 2

The second example illustrates a case with more complex dimensionality and higher noise levels. It corresponds to site b23, from the Betic Chain magnetotelluric dataset. The variances in the magnetotelluric tensor components, with the non-diagonals slightly larger than the diagonals, correspond approximately to 1% noise for the shortest periods, and increase up to 30% for the longest ones.

Comparison of the three approaches

The invariants and their errors were estimated using approaches **a**, **b** and **c**. In approaches **b** and **c** the values of n and N were the same as in COPROD2: n=1000 in approach **b** and N=400 (from an initial sample with n=100) in approach **c**.

Employing the three approaches, the values of I_3, I_4, I_5 and I_6 (Figure 3.3) are comparable and the errors resemble those of the components of the magnetotelluric tensor, which increase with the period. However, at long periods, biases between the values computed from the different approaches are appreciable and the statistical errors (approaches **b** and **c**) present extremely large values which were not observed in the COPROD2 example, due to the low noise level in the MT tensor components in this long period band.

I_7 and Q have the same behaviour as in the first example, with the small values and error bars of Q and the extremely large error bars and biases of I_7. Between 3s and 30s, as Q becomes large, I_7 becomes more precise, as in site 85_314. Hence, the estimation of I_7 depends not only on the noise level, but also on the value of Q.

3.2.4 Concluding remarks on error estimation

After these comparisons, having obtained similar estimations of the invariants and their errors, it can indeed be concluded that the three approaches can be used to estimate the invariants values and their errors, except for cases with large error levels or small values of Q. However, in order to avoid the possible biases that appear with approaches **b** and **c**, the use of classical error propagation (approach **a**) is preferred.

A similar test was performed for the estimation of the strike directions and distortion parameters. In these cases, errors are better estimated using random noise generation (approach **b**). Using classical error propagation, since these errors depend on the derivatives of trigonometric functions, the error bars would be considerably large, especially for high noise levels in the MT tensor components.
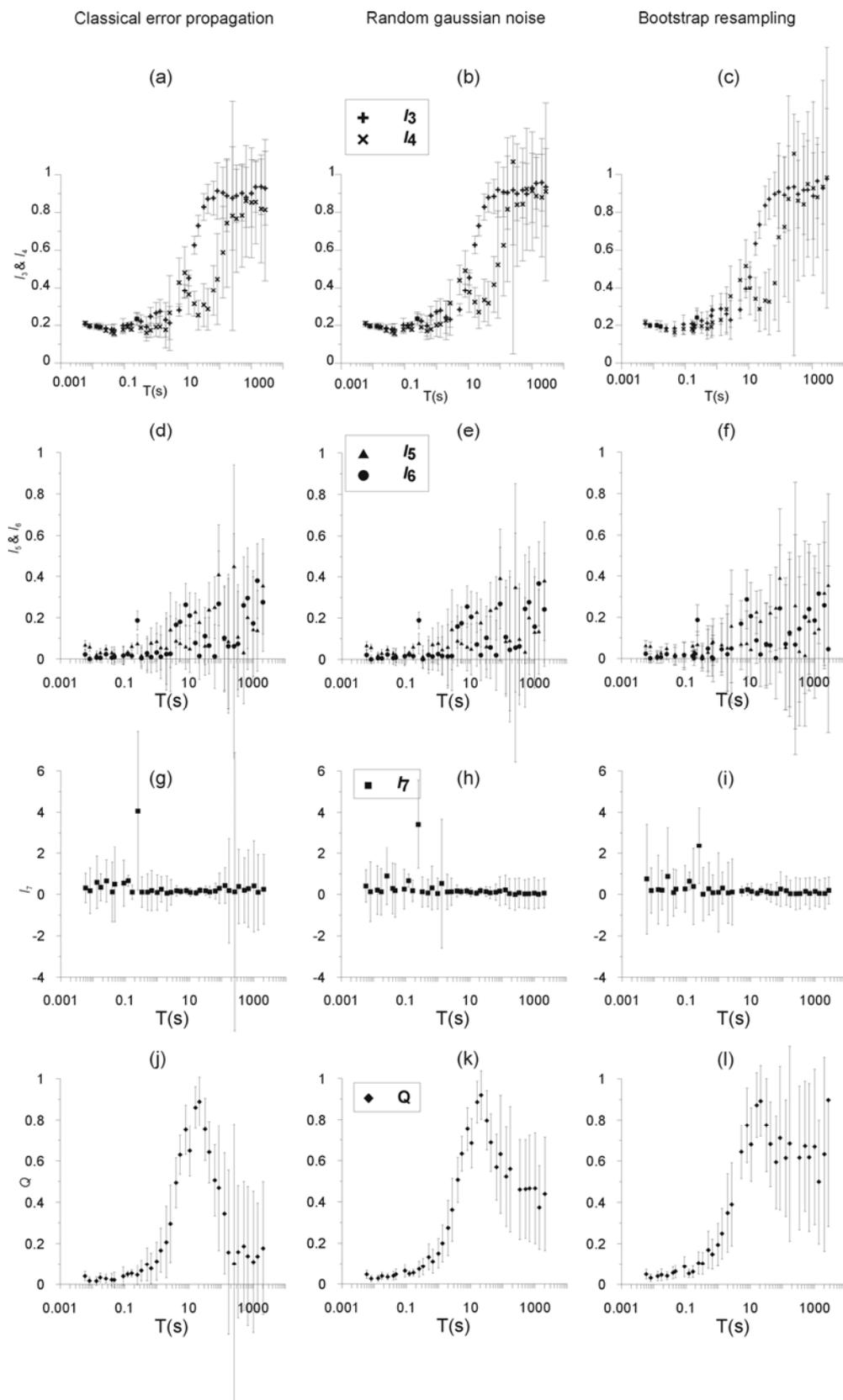
Figure 3.3: Invariant values corresponding to site b23 from the Betic dataset and their errors computed using the three approaches. For random Gaussian Noise, n=1000. In the bootstrap 400 realizations from an original sample with 100 elements for each invariant were used. I_3 and I_4 (a), (b) and (c); I_5 and I_6 (d), (e) and (f), I_7 (g), (h) and (i) and Q (j), (k) and (l).

3.3. The Threshold Value

For the determination of the dimensionality using WAL invariants, it is necessary to decide if an invariant can be considered null or not. This is a compromise between the threshold value (τ hereafter) and the error bar of each invariant. After considering the possible relationships between the thresholds and the confidence intervals of the invariants, the following criteria were adopted:

1) Ascertaining whether I_7 is undefined or not, by observing the values of Q, which is controlled by a threshold τ_Q (below which Q is regarded as too small a quotient of I_7), and the value of I_7:

    1.a) $Q < \tau_Q$ and/or $I_7 > 1 \Rightarrow I_7$ is undefined,

    1.b) $Q > \tau_Q$ and $I_7 < \tau \Rightarrow I_7 \approx 0$,

    1.c) $Q > \tau_Q$ and $I_7 > \tau \Rightarrow I_7 \neq 0$.

Given that I_7 is highly sensitive to the data errors, its validity has been limited to its dependence on the value of Q. Otherwise, the errors of I_7 would dominate the dimensionality estimation, which would be always undetermined.

2) The rest of the invariants, I_3 to I_6, are considered null or not depending on the possible values of I_k and σ_k (Figure 3.4):

    2.a) $I_k - \sigma_k$ and $I_k + \sigma_k < \tau \Rightarrow I_k \approx 0$ (Figure 3.4a),

    2.b) $I_k > \tau$ and $I_k - \sigma_k > \tau \Rightarrow I_k \neq 0$ (Figure 3.4b),

    2.c) $I_k < \tau$ and $I_k + \sigma_k > \tau \Rightarrow I_k \neq 0$ (Figure 3.4c), (Note that I_k would be regarded null had the error not been taken into account),

    2.d) $I_k > \tau$ and $I_k - \sigma_k < \tau \Rightarrow I_k \neq 0$ (Figure 3.4d),

    2.e) $I_k + \sigma_k > 1 \Rightarrow I_k$ is undefined (Figure 3.4f),

which can be reduced to the conditions:

$$I_k + \sigma_k < \tau \Rightarrow I_k \approx 0, \qquad (3.9)$$

$$\tau < I_k + \sigma_k \leq 1 \Rightarrow I_k \neq 0, \qquad (3.10)$$

and I_k is undefined for the rest of the cases.

If any of the invariants I_3 to I_6 are undefined (Figure 3.4e) the dimensionality of the corresponding tensor cannot be determined.
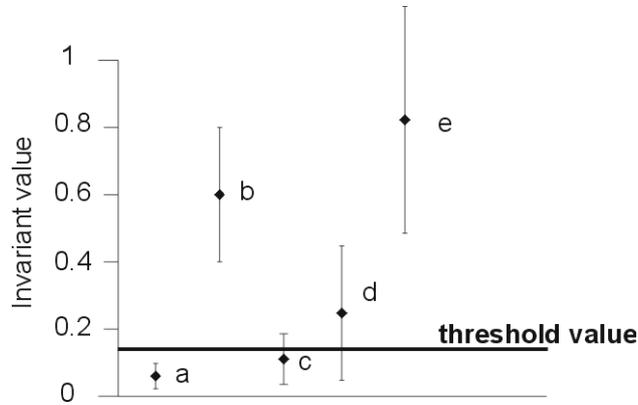
Figure 3.4: Different possibilities for the invariant values and their errors (τ= threshold value). (a) $I_k - \sigma_k < \tau$ and $I_k + \sigma_k < \tau$; (b) $I_k - \sigma_k > \tau$ and $I_k + \sigma_k > \tau$; (c) $I_k < \tau$ and $I_k + \sigma_k > \tau$; (d) $I_k > \tau$ and $I_k - \sigma_k < \tau$ and (e) $I_k + \sigma_k > 1$. Only in case (a) is invariant I_k considered to be zero.

The choice of the threshold value τ is a subjective decision, but it is important to have an appropriate range of values that works for a correct dimensionality analysis. A high value of τ produces invariants that are considered null, resulting in too simple a structure (1D for all sites and periods). On the contrary, a small value of τ implies that the invariants cannot be considered null and it gives rise to a more complex structure (3D in general).

An exhaustive study was carried out with the aim of determining the optimum range of threshold values that provides a stable and consistent dimensionality pattern, taking into account the errors of the invariants. This task was performed using the examples from the previous section, site 85_314 from COPROD2 and site b23 from the Betics, plus two additional Betics sites.

In both examples, the WAL dimensionality analysis was performed for different threshold values and the results were compared to determine those most appropriate. The threshold values tested were 0.08, 0.10, 0.12, 0.15, 0.20 and 0.30. In the second example, using three sites from the Betics dataset, a more detailed study was done, including an analysis on the determination of the strike directions and distortion parameters, and a comparison with decomposition techniques.

With regard to the threshold for invariant Q, τ_Q, which conditions the determination of invariant I_7, its value was fixed to 0.10. Posterior tests changing the value of τ_Q resulted in small

changes in the dimensionality. Nevertheless, I recommend not using a τ_Q value lower than 0.1, as the value of I_7 would be considered undefined even if it was much smaller than unity.

The dimensionality pattern corresponding to all the registered periods from site 85_314, for the different threshold values, is illustrated in Figure 3.5. Those cases undetermined as a consequence of the errors are also included, the number of which decreases as the threshold value increases. In general, and apart from the extreme values, 0.08 and 0.3, the dimensionality is stable among different threshold values, especially for the longest periods. Some particular changes are observed between the intermediate values, which are a consequence of the error bars crossing the threshold values. The same dimensionality analysis but, without considering errors, would result in the same dimensionality pattern, although without any undetermined cases.
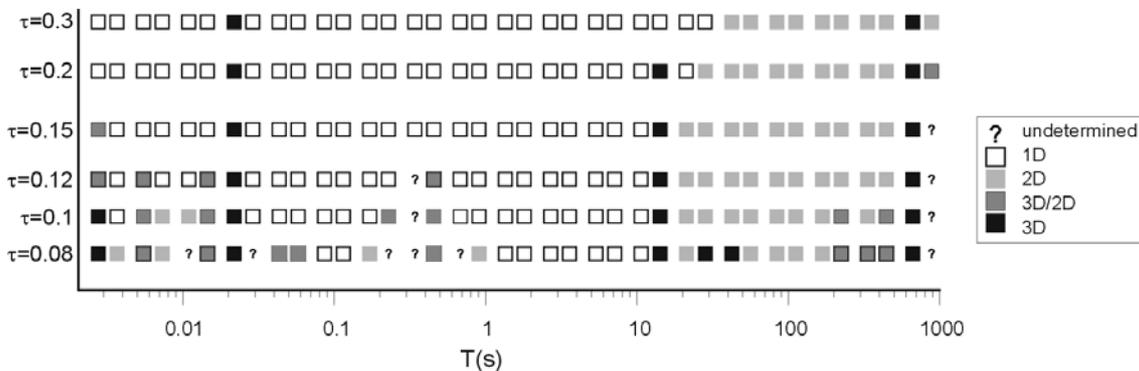
Figure 3.5: Dimensionality cases for site 85_314 from the COPROD2 dataset using different threshold values: 0.08, 0.10, 0.12, 0.15, 0.20 and 0.30, considering data errors.

Figure 3.6 shows the dimensionality cases obtained for each threshold value for the three sites from the Betics dataset, b01, b23 and b40. Independent of the threshold value, dimensionality remains undetermined for those periods with relative errors in the MT tensor components greater than approximately 30%.

The extreme values τ=0.08, τ=0.20 and τ=0.30 provide dimensionality patterns which are inconsistent. On the contrary, intermediate values, τ = 0.1, 0.12 and 0.15, give a quite stable pattern and a characteristic behaviour can be observed at each site. Site b01 reflects a 3D/2D dimensionality for some periods up to 0.12s using τ = 0.1, 0.12 and 0.15. Site b23 appears as 2D up to 0.2 s, for any threshold value, and as 3D/2D for longer periods up to 50 s, using τ = 0.12, τ = 0.15 and even τ = 0.2. Site b40 shows a 1D dimensionality at short periods (up to 1 s) for the three intermediate τ values.
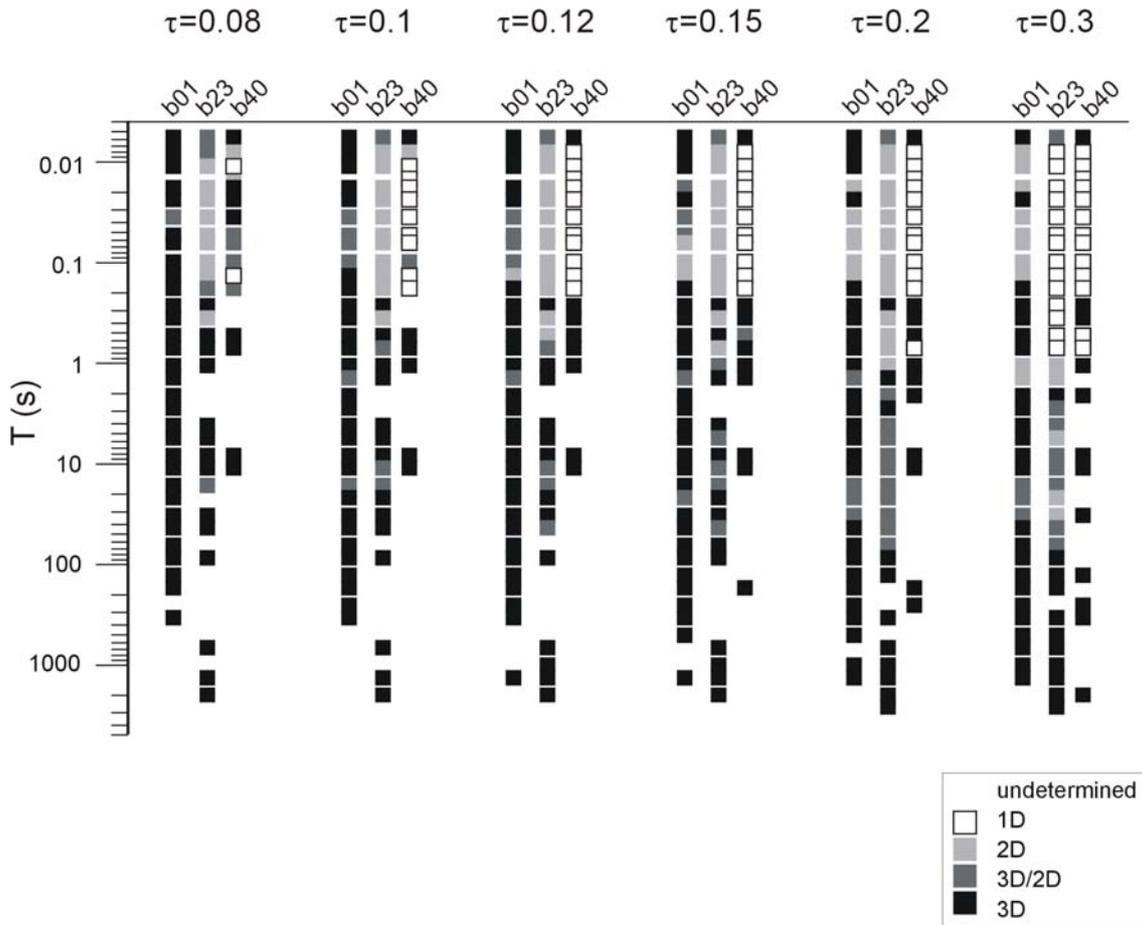
Figure 3.6: Dimensionality cases for sites b01, b23 and b40 from the Betics MT dataset using different threshold values: 0.08, 0.10, 0.12, 0.15, 0.20 and 0.30.

In order to confirm whether this description is valid or not, another test was performed, considering the determination of the strike angles and the distortion parameters. For all the periods of these three sites, two types of analysis were performed:

- Analysis A: the strike directions and errors corresponding to a 2D structure were determined (θ_1 and θ_2 in equation 2.19, equivalent to Swift's angle, equation 2.5).

- Analysis B: a 3D/2D structure was assumed, computing the strike and distortion angles and their errors (θ_3 in equation 2.23 and φ_t and φ_e in equations 2.24 and 2.25, equivalent to Groom and Bailey decomposition).

The strikes and errors obtained using both analyses made it possible to constrain frequencies for which a strike direction and/or distortion parameters could be determined. If θ_1 and θ_2 are similar with constant values and small error bars, the structure can be considered 2D.

If θ_3 is the angle with small error bars and constant values of distortion angles, it can be considered 3D/2D. For the analysed sites the main results are:

Site b01 (Figure 3.7): Analysis A (Figure 3.7a) gives a good determination of the strike directions up to 200 s, where the errors in θ_1 and θ_2 become large. However, since values θ_1 and θ_2 are dissimilar, it cannot correspond to a 2D structure. Analysis B (Figure 3.7b) shows large error bars, except for the period range between 0.02 s and 0.2 s, which can be described as 3D/2D because θ_3 and the distortion angles have a constant value.
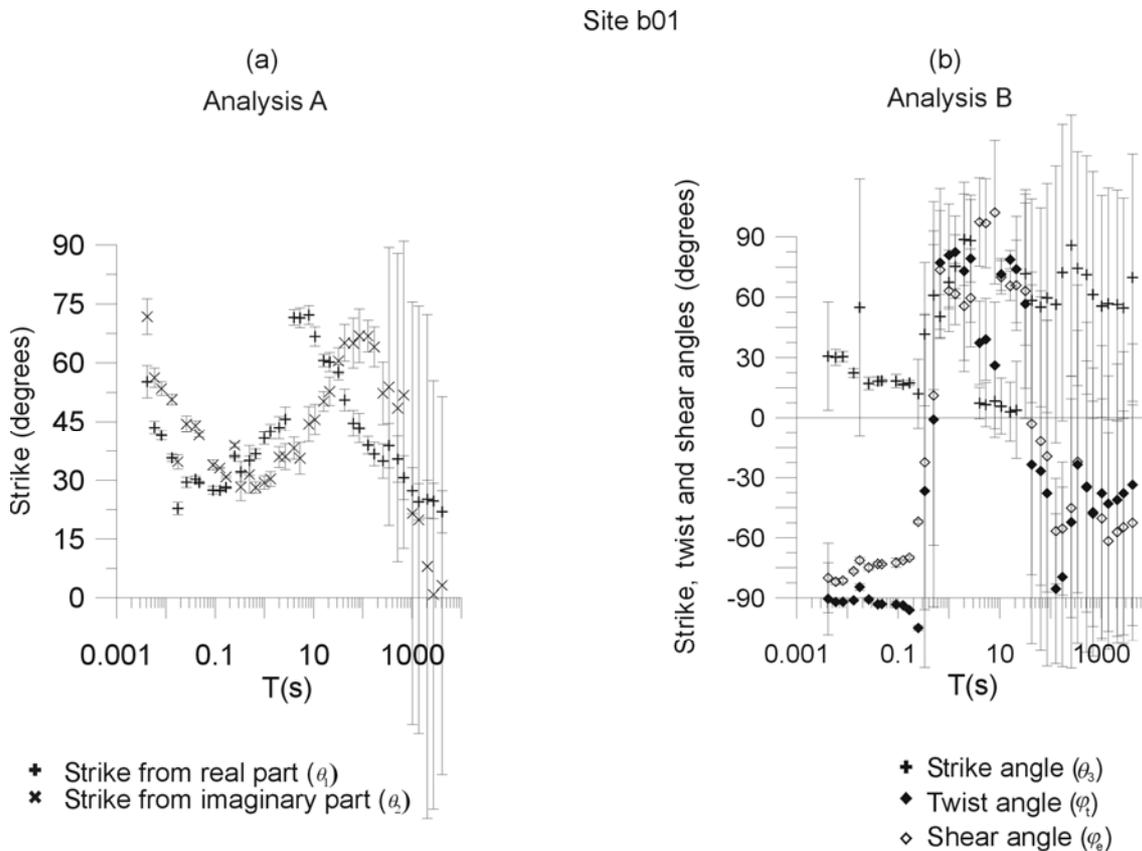
Figure 3.7: Strike directions and distortion parameters for site b01: (a) analysis A, (b) analysis B.

Site b23: Analysis A gives a good determination of the strike for the lowest periods up to 0.2s, which have the same value $\theta_1 = \theta_2$ (Figure 3.8a). Analysis B shows that constant strike direction and twist and shear angles can be inferred between 5 s and 50 s (Figure 3.8b).
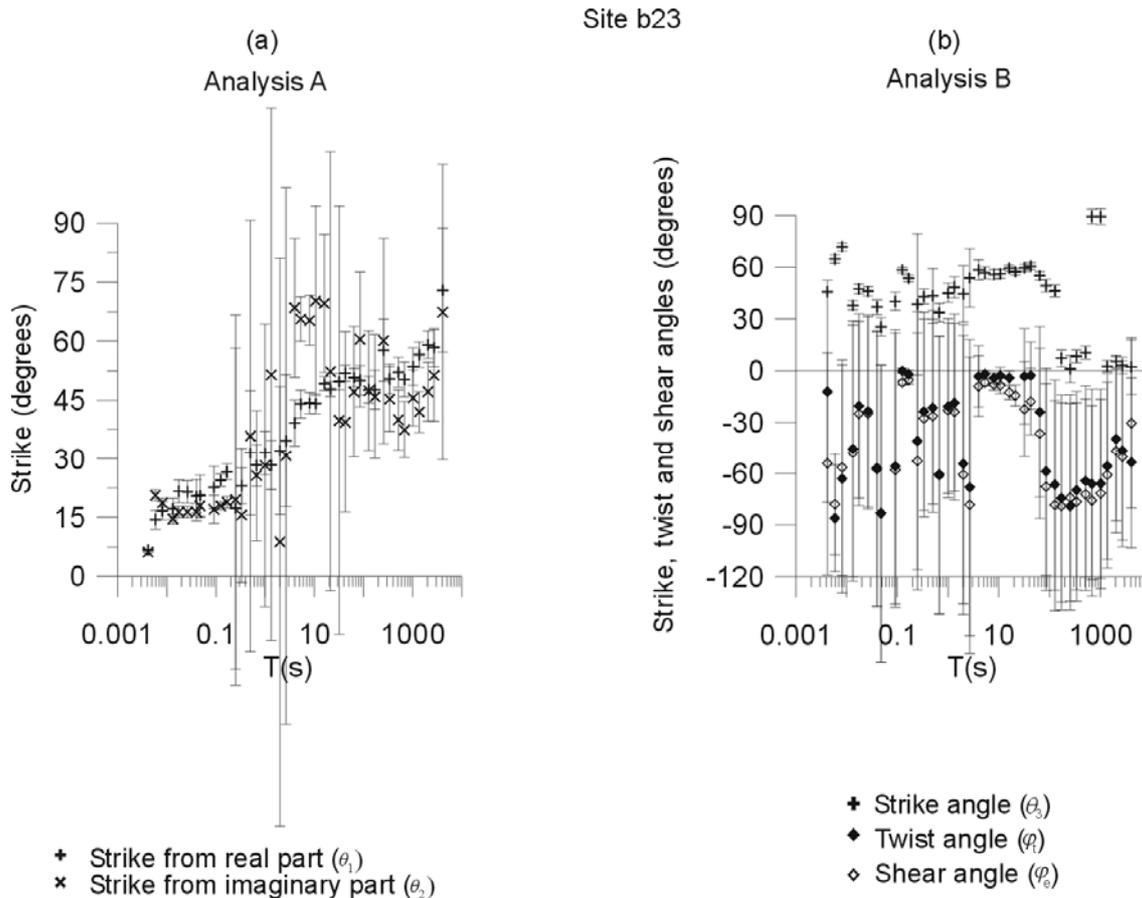
Figure 3.8: Strike directions and distortion parameters for site b23: (a) analysis A, (b) analysis B.

Site b40: neither analyses A and B nor the computation of θ_1, θ_2 and θ_3 (Figures 3.9a and 3.9b) give good determination of a possible 2D or 3D/2D structure. However, up to periods of about 1s, the MT tensor corresponds to a 1D case ($M_{xy} = -M_{yx}$ and $M_{xx} = M_{yy} = 0$), and the computed apparent resistivities (Figure 3.9c) and phases (Figure 3.9d) have the same value for xy and yx modes. These values are the same as those obtained from invariants I_1 and I_2 (equations 2.23 and 2.24). This is consistent with a 1D interpretation of the data for the short and middle periods of this site.

The results from these analyses correspond to the dimensionality obtained using threshold values between 0.1 and 0.15. Hence, this threshold range is appropriate for a dimensionality analysis of data with up to 30% noise in the impedance tensor components.
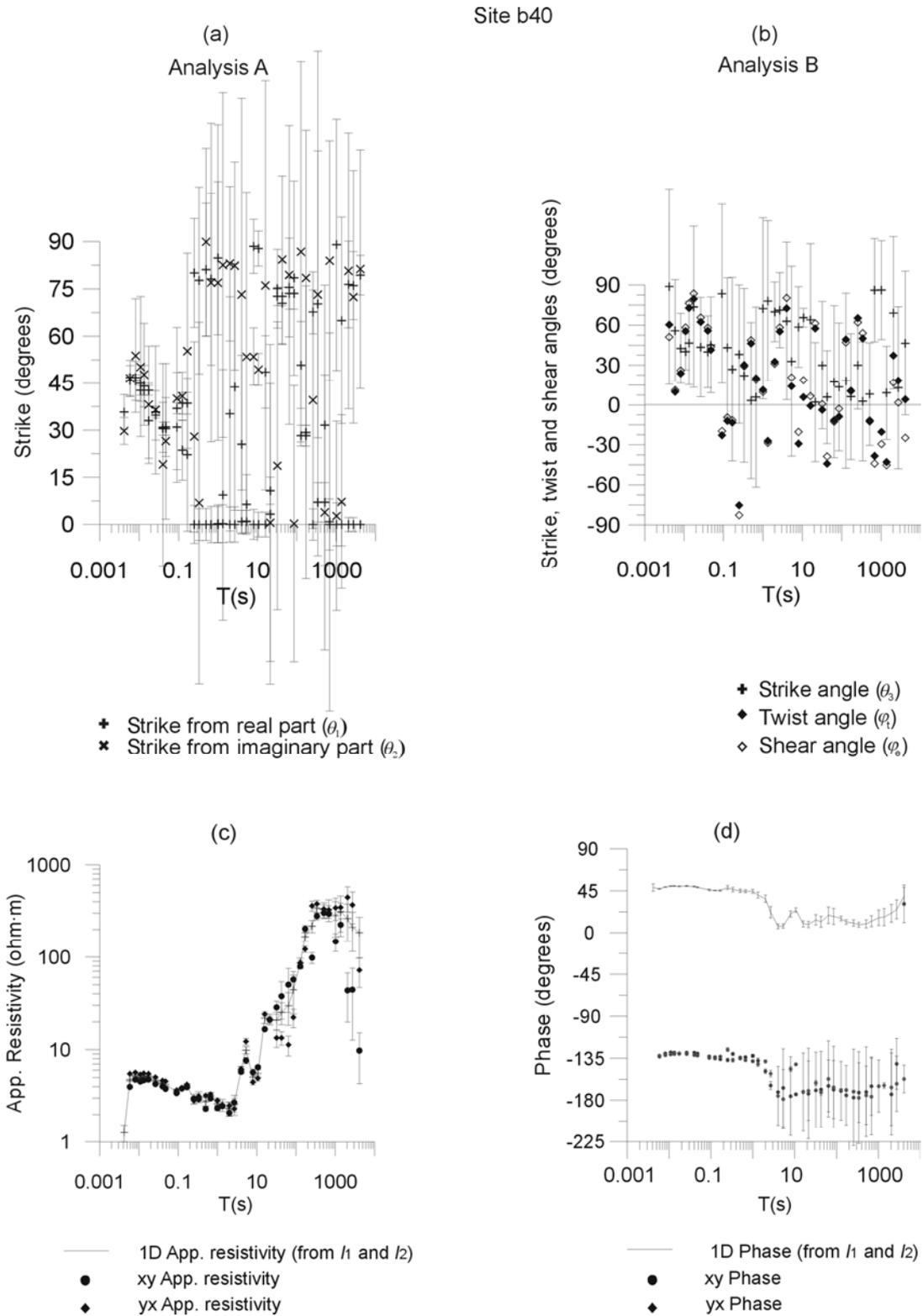
Figure 3.9: Strike directions and distortion parameters for site b40: (a) analysis A, (b) analysis B, (c) and (d): xy and yx apparent resistivities and phases computed directly from the MT tensor and I_1 and I_2 invariants.

3.4. Practical Criteria

After performing these tests on the error treatment and the threshold value, these final recommendations are given:

(1) Determine the errors of the invariants using any of the described approaches, **a**, **b** or **c**. However, classical error propagation (approach **a**) is recommended to avoid biases.

(2) Determine the dimensionality using WAL criteria with a threshold value between 0.1 and 0.15, letting I_7 be undetermined if Q is below τ_Q, and giving consideration to the error bars. Note that dimensionality will be well determined when relative errors in M are not greater than approximately 30%.

(3) Compute the strike directions and/or distortion angles corresponding to 2D and 3D/2D cases and their errors, using Random Gaussian Noise generation.

3.5. WALDIM Program

A Compaq Visual Fortran 2000® application, termed WALDIM, was created. It is a tool to automatically obtain the dimensionality analysis from a set of raw or synthetic MT data, based on WAL criteria while considering noise in the data.

The main functions of this program are to compute WAL invariants, related parameters and errors corresponding to each MT tensor and to determine the dimensionality, following WAL criteria, and according to the errors and the threshold value. In those cases related to two-dimensional structures and/or the presence of distortion, apart from the dimensionality, the strike and distortion parameters with their errors are provided as relevant information.

The strike angles are all determined in the first quadrant. As stated in chapter 1, the 90º ambiguity should be solved using the tipper information.

The program also solves some inconsistencies that can appear regarding two-dimensionality and strike directions: On some occasions, WAL criteria indicate 2D dimensionality, whereas θ_1 and θ_2 strike angles have significantly different values. This is a consequence of having defined a lower than required threshold value. The program solves this by changing the dimensionality of these cases into 3D/2D and assigning θ_3, ϕ_1 and ϕ_2 as the strike direction and distortion parameters respectively. The difference between θ_1 and θ_2 is set to

a maximum of 10°, or a minimum of 80°, given that one of the angles can have a 90° ambiguity, although these values may be modified.

A broad range of parameters and options remain open for the user, which may be chosen depending on the quality or type of data, or to continue investigating the threshold values and error analysis. The values of thresholds τ and τ_Q can be chosen by the user. Data errors can be those in the MT tensor components, or can be computed as a noise level specified by the user. The error estimation in the computed parameters can be done as in approach **a** (classical error propagation) or as in **b** [1] (Random Gaussian Noise generation), which in the program is referred to as "statistical approach". Bootstrap resampling can be performed by calling an external program.

Another option of WALDIM is to classify the dimensionality into groups of periods for each site such as to have a more general image of the dimensionality therein. In this classification, a specific period range is divided into groups made up of multiples or fractions of the decades contained within this range. For each site, the dimensionality of a group is the mode of the data in the group. In the case that the mode has more than one dimensionality type, priority is given to that which is less complex. If the dimensionality of the group needs computation of strike and distortion parameters, these and their errors are the average and standard deviation of the group data with that type of dimensionality. If inconsistencies in the determination of 2D strike angles appear, these are solved by changing the cases to 3D/2D as previously explained.

Figure 3.10 schematizes the general flowchart of the program. It has been tested for different datasets, some of which appear in the dimensionality studies carried out in this thesis.

3.6. Conclusions

A method has been developed to determine the geoelectric dimensionality using the WAL invariants, taking into account the data errors. The errors of the invariants and related parameters were estimated using three approaches: classical error propagation, generation of Random Gaussian Noise and bootstrap resampling. Different threshold values were tested to ensure a stable dimensionality pattern.

The errors of the invariants can be properly estimated by classical error propagation, but the generation of random values is more robust, thus ensuring stability in the errors of strike direction and distortion parameters. The use of a threshold value between 0.1 and 0.15 is recommended for real data of medium to high quality.

---

[1] Subroutine Gasdev (Press *et al.*, 1992)

A program named WALDIM to perform MT dimensionality analysis using WAL criteria and considering data noise, has been developed. It offers several options with regard to error estimation and threshold values, and has been successfully tested.
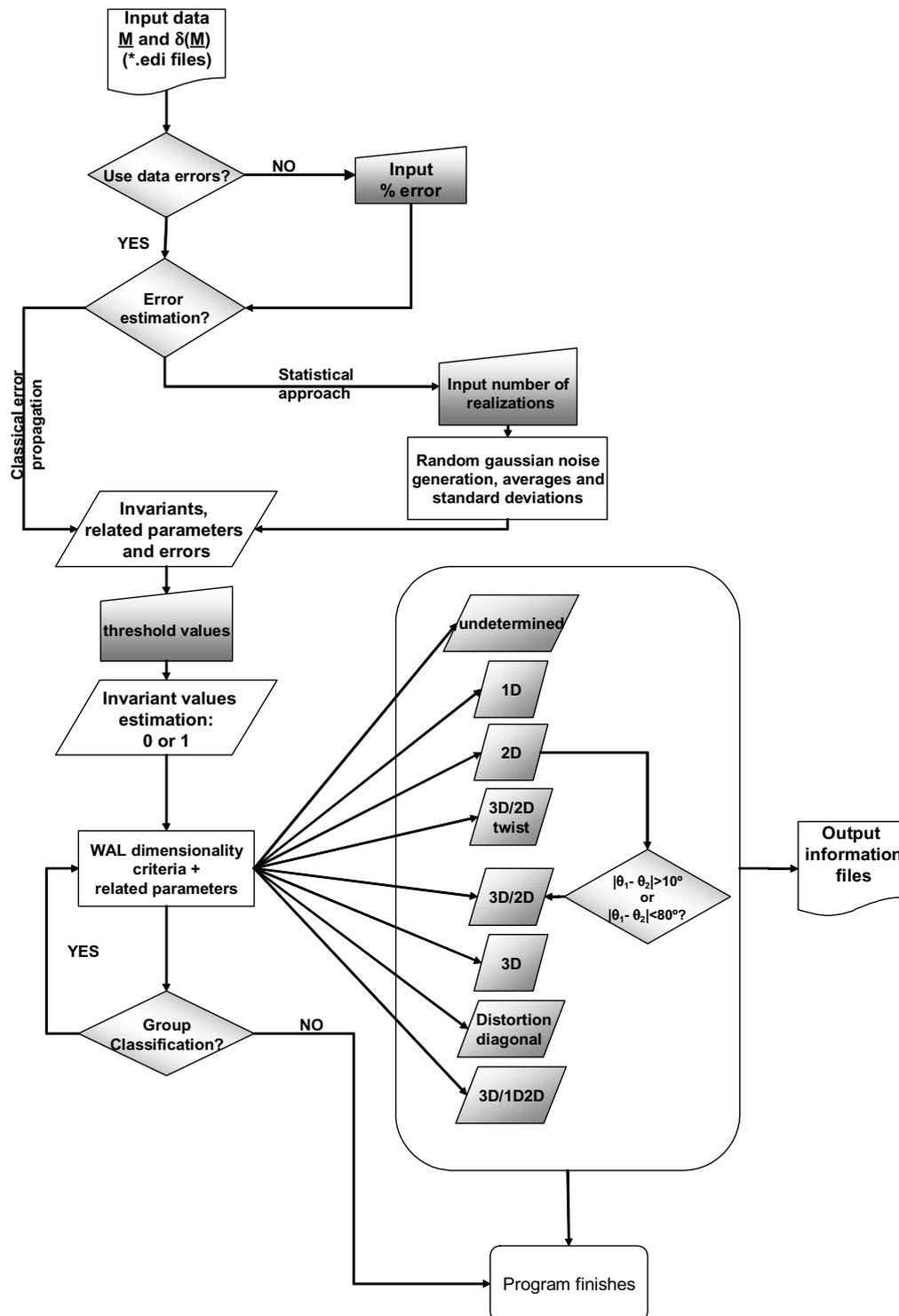
Figure 3.10: General overview flowchart of WALDIM program.

Chapter 4: Improving Bahr's Invariant Parameters Using the WAL Approach

This chapter presents a comparison between Bahr and WAL parameters and their dimensionality criteria. It led to a proposal of a unification of the existing threshold values used in these dimensionality analyses. And, to the improving of Bahr's criteria with the addition of the Q parameter, from the WAL invariants set. The new method was hence termed the Bahr-Q method.

The motivation for this work came from the previous work of Ledo *et al.* (2002b), on the limitations of a 2D interpretation of 3D data. Among different aspects concerning their effects in modelisation, it demonstrated the misuse of some of the Bahr parameters at ascertaining whether data are 3D or not.

Part of this work has been already published in Martí *et al.* (2005).

4.1 Introduction

As already described in chapter 2, both the Bahr and WAL methods use a set of rotational invariant parameters of the magnetotelluric tensor and establish a classification of these values to determine the kind of dimensionality associated with the measured data.

Among the four parameters defined by Bahr (1991) (Chapter 2, section 2.4), only the skew, κ, and regional skew, η, are commonly used to test the validity of dimensionality, as opposed to the four of them being fully taken into account to characterise it. The use of the

WAL method (Chapter 2, section 2.5), although limited, is increasing (Jones *et al.*, 2002; Martí *et al.*, 2004).

When the data are of good quality (low noise and low distortion) and the subsurface materials being imaged can be approximated by 2D structures, both methods work reasonably well and give similar results. When dealing with more complex structures and with real data, propagation of data errors and oversimplification of the physical models used will lead, in most cases, to more confusing situations.

The aim of this work was to study the Bahr and WAL methods and to propose a new one that makes both dimensionality methods consistent.

To accomplish this, the analytical relationships between both sets of parameters were derived and the threshold values of each method were revised and compared. To illustrate the results of this study one set of synthetic data and two sets of real data were used, including error effects.

4.2 Bahr and WAL Methods

The dimensionality information given by both sets of parameters, and the recommended threshold values are summarised in Table 4.1.

Case	Bahr invariants	WAL invariants	Dimensionality case
1	$\kappa < 0.1, \Sigma < 0.1$	$I_K=0$ (k=3-7)	1D
2	$\kappa < 0.1, \Sigma > 0.1$	$I_3 \neq 0$ or $I_4 \neq 0$	2D
3	$\kappa > 0.1, \mu < 0.05$		3D/1D
3		$I_3 \neq 0$ or $I_4 \neq 0$ and $I_5 \neq 0$	a) 3D/2D (twist) (if I_7=0) b) 3D/2D or 3D/1D (if I_7 undefined)
4	$\kappa > 0.1, \mu > 0.05, \eta < 0.1$	$I_6 \neq 0$	3D/2D
5	$\kappa > 0.1, \mu > 0.05, \eta > 0.3$	$I_7 \neq 0$	3D
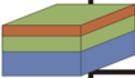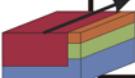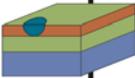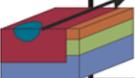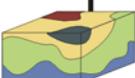
Table 4.1: Dimensionality cases according to Bahr and WAL rotational invariants. In case 3, red and green colours refer to Bahr and WAL criteria respectively. The WAL method can distinguish cases 3a) from 3b) whereas the Bahr method cannot.

One of the main limitations of the Bahr method is that the threshold values do not have a justified physical or mathematical meaning, nor are they set in a statistical framework. As is shown below, the use of only these four parameters is insufficient to completely characterise the dimensionality.

The number of WAL invariants used for determining dimensionality is five (I_3 to I_7) plus invariant Q, which determines whether I_7 is meaningless or not. This total of six can be reduced to five since I_3 and I_4 are used together in the dimensionality classification.

In order to compare the reliability of both methods, an example was used which consisted of a magnetotelluric tensor $\underline{M}$, clearly representing a two-dimensional structure with different phases in the xy and yx polarisations. Table 4.2 shows the components of this tensor and the dimensionality interpretation using Bahr and WAL criteria. It is clear that in some circumstances, Bahr's criteria may lead to incorrect interpretations.

The use of only four parameters in the Bahr method is an important limitation in determining dimensionality. Furthermore, the classification of the dimensionality types, as well as the threshold values, should be revised such that they are concise and consistent.

As a first step, the next section presents the analytic relationships between the Bahr and WAL parameters.

MT Tensor	Bahr invariants	WAL invariants	
$M=\begin{bmatrix} 0 & 25+9i \\ -15-12i & 0 \end{bmatrix}$	$\kappa = 0$ $\mu = 0$ $\eta = 0$ $\Sigma = 0.05$	$I_1=20$ $I_2=10.5$ $I_3=0.25$ $I_4=0.14$ $I_5=0$ $I_6=0$ $I_7=0$ $Q=0.39$	
	κ and $\Sigma < 0.1$, η and $\mu \approx 0 \rightarrow$ 1D or 3D/1D	I_3 and $I_4 \neq 0$, I_5, I_6 and $I_7 = 0 \rightarrow$ 2D	

Table 4.2: Example showing how both methods, using the threshold values established by the authors, lead to different interpretations.

4.3 Analytical Relationships and Equivalencies for Ideal Cases

Using the same decomposition as in Weaver *et al.* (2000), where the magnetotelluric tensor components are expressed as a function of the complex parameters $\zeta_j=\xi_j+i\cdot\eta_j$ (eq. 2.9), Bahr rotational invariants were rewritten as:

$$\kappa = \frac{\left(\xi_1^2 + \eta_1^2\right)^{1/2}}{\left(\xi_4^2 + \eta_4^2\right)^{1/2}}, \quad \text{(Swift Skew)} \tag{4.1a}$$

$$\mu = \frac{\left(\left|\xi_3\eta_2 - \xi_2\eta_3\right| + \left|\xi_1\eta_4 - \xi_4\eta_1\right|\right)^{1/2}}{\left(\xi_4^2 + \eta_4^2\right)^{1/2}}, \tag{4.1b}$$

$$\eta = \frac{\left(\left|\xi_3\eta_2 - \xi_2\eta_3 - \xi_1\eta_4 + \xi_4\eta_1\right|\right)^{1/2}}{\left(\xi_4^2 + \eta_4^2\right)^{1/2}}, \quad \text{(Regional skew or Phase sensitive skew)} \tag{4.1c}$$

$$\Sigma = \frac{\xi_2^2 + \xi_3^2 + \eta_2^2 + \eta_3^2}{\xi_4^2 + \eta_4^2}. \tag{4.1d}$$

Departing from these expressions, the analytical relationships were obtained by expressing Bahr parameters as functions of WAL invariants and are given as:

$$\kappa = f(r,I_5,I_6) = \sqrt{\frac{1+r^2}{1-\frac{(I_5-I_6)^2}{2\left(1-I_5I_6+\sqrt{1+I_5^2I_6^2-I_5^2-I_6^2}\right)}+\frac{r^2}{2}\left(1-I_5I_6+\sqrt{1+I_5^2I_6^2-I_5^2-I_6^2}\right)}-1} = \sqrt{\frac{1+r^2}{1-\frac{a}{2b}+\frac{r^2b}{2}}-1} \tag{4.2a}$$

$$\mu = f(r,I_5,I_6,I_7,Q) = \sqrt{\frac{|I_7Q+I_6|+|I_6|}{\frac{1}{r}\left(1-\frac{(I_5-I_6)^2}{2\left(1-I_5I_6+\sqrt{1+I_5^2I_6^2-I_5^2-I_6^2}\right)}\right)+\frac{r}{2}\left(1-I_5I_6+\sqrt{1+I_5^2I_6^2-I_5^2-I_6^2}\right)}} = \sqrt{\frac{|I_7Q+I_6|+|I_6|}{\frac{1}{r}\left(1-\frac{a}{2b}\right)+\frac{rb}{2}}} \tag{4.2b}$$

$$\eta = f(r,I_5,I_6,I_7,Q) = \sqrt{\frac{|I_7Q|}{\frac{1}{r}\left(1-\frac{(I_5-I_6)^2}{2\left(1-I_5I_6+\sqrt{1+I_5^2I_6^2-I_5^2-I_6^2}\right)}\right)+\frac{r}{2}\left(1-I_5I_6+\sqrt{1+I_5^2I_6^2-I_5^2-I_6^2}\right)}} \tag{4.2c}$$

$$\Sigma = f(r,I_3,I_4,I_5,I_6) = \frac{I_3^2+I_4^2r^2}{1-\frac{(I_5-I_6)^2}{2\left(1-I_5I_6+\sqrt{1+I_5^2I_6^2-I_5^2-I_6^2}\right)}+\frac{r^2}{2}\left(1-I_5I_6+\sqrt{1+I_5^2I_6^2-I_5^2-I_6^2}\right)} = \frac{I_3^2+I_4^2r^2}{1-\frac{a}{2b}+\frac{r^2b}{2}} \tag{4.2d}$$

where a and b depend on I_5 and I_6: $a = (I_5 - I_6)^2$ and $b = 1 - I_5 I_6 + \sqrt{1 + I_5^2 I_6^2 - I_5^2 - I_6^2}$. r is the quotient between I_2 and I_1, which appears in all the relationships due to the different normalisation used in each set of parameters.

Relating these identities to WAL dimensionality criteria, parameter κ is the one that distinguishes between undistorted and distorted cases (depending on whether I_5 and I_6 are null or not). μ and η also depend on the product $I_7 \cdot Q$, but not on their individual values. Σ is the only parameter that depends on I_3 and I_4, and, consequently, that deals with two-dimensionality.

The analytical relationships presented allow the particular expressions of Bahr parameters to be easily obtained for the ideal cases following the WAL conditions (in parenthesis):

1) 1D ($I_3 - I_6$=0 and I_7=0 or Q=0)

$$\kappa = \mu = \eta = \Sigma = 0 . \tag{4.3}$$

2) 2D (I$_3$ or $I_4 \neq 0$, I_5, I_6=0 and I_7=0 or Q=0)

$$\kappa = \mu = \eta = 0 , \tag{4.4a}$$

$$\Sigma = f_4(r, I_3, I_4) = \frac{I_3^2 + I_4^2 r^2}{1 + r^2} . \tag{4.4b}$$

3) a) 3D/2Dtwist and b) 3D/1D2D (I_3 or $I_4 \neq 0$, $I_5 \neq 0$, I_6=0 and I_7=0 or Q=0)

$$\mu = \eta = 0 , \tag{4.5a}$$

$$\kappa = f_1(r, I_5) , \tag{4.5b}$$

$$\Sigma = f_4(r, I_3, I_4, I_5) . \tag{4.5c}$$

Cases 3a and 3b are non-distinguishable using the Bahr method because it is not possible to know which of I_7 or Q is null.

4) 3D/2D ($I_6 \neq 0$ and I_7=0): κ, μ and Σ have the values corresponding to the general expressions (eqs. 4.2a-4.2d):

$$\eta = 0 , \tag{4.6a}$$

$$\kappa = f_1\left(r, I_5, I_6\right),$$ (4.6b)

$$\mu = f_2\left(r, I_5, I_6\right),$$ (4.6c)

$$\Sigma = f_4\left(r, I_3, I_4, I_5, I_6\right).$$ (4.6d)

5) 3D ($I_7 \neq 0$ and $Q \neq 0$): κ, μ, η and Σ have the values corresponding to the general expressions 4.2a-4.2d. However, there is no condition on the values of $I_3 - I_6$, therefore these invariants could have any value from 0 to 1.

$$\kappa = f_1\left(r, I_5, I_6\right),$$ (4.7a)

$$\mu = f_2\left(r, I_5, I_6, I_7, Q\right),$$ (4.7b)

$$\eta = f_3\left(r, I_5, I_6, I_7, Q\right),$$ (4.7c)

$$\Sigma = f_4\left(r, I_3, I_4, I_5, I_6\right).$$ (4.7d)

From these relationships some important points arise that modify the Bahr criteria (Table 4.1):

a) If the dimensionality is 1D or 2D, parameter μ must be null.

b) $\kappa \neq 0$ and $\mu \neq 0$ conditions are not necessary for three-dimensionality.

c) If the structure is not 3D, η vanishes.

4.4 Threshold Values

Given that in real situations data are affected by noise and that geoelectric structures do not exactly fit the assumed ideal cases, invariant values are never precisely zero. Weaver *et al.* (2000) address this problem by introducing a threshold value (τ_W), the same for all invariants I_3 to I_7, beneath which they are considered to be zero.

In order to make the Bahr and WAL criteria equivalent, first it was necessary to obtain the relationships between WAL and Bahr threshold values and then to redefine the thresholds used in the Bahr method. It was accomplished with two approaches: an analytical development and a statistical one.

a) Analytical Development

As a first approach, Bahr thresholds were computed from the analytical relationships, using the WAL threshold τ_W in place of invariants I_3 - I_7. The chosen value of Q was unity, thus the product $I_7 \cdot Q$ was replaced by τ_W. Parameters a and b in equations 4.2a to 4.2d were approximated by $a=(2\cdot\tau_W)^2$ and $b=2$, after replacing I_5 and I_6 by the threshold value τ_W, excluding second order terms.

It was assumed that all WAL invariants are equal to the threshold value. This is not the situation for every dimensionality case but makes it possible to obtain preliminary expressions of Bahr thresholds. τ_κ, τ_μ, τ_η and τ_Σ depend on τ_W and r (the relationship between I_2 and I_1):

$$\tau_\kappa = \sqrt{\frac{1+r^2}{1+r^2-\tau_W^2}-1}\,, \tag{4.8a}$$

$$\tau_\mu = \sqrt{\frac{3\tau_W}{\sqrt{\frac{1}{r}\left(1-\tau_W^2\right)+1}}}\,, \tag{4.8b}$$

$$\tau_\eta = \sqrt{\frac{2\tau_W}{\sqrt{\frac{1}{r}\left(1-\tau_W^2\right)+1}}}\,, \tag{4.8c}$$

$$\tau_\Sigma = \frac{\tau_W^2(1+r^2)}{1+r^2-\tau_W^2}\,. \tag{4.8d}$$

Despite the approximation used, these expressions are useful to see which of the Bahr parameters are the most sensitive to τ_W and r. The dependencies of Bahr thresholds on the WAL threshold, for different values of r, are represented in Figure 4.1.
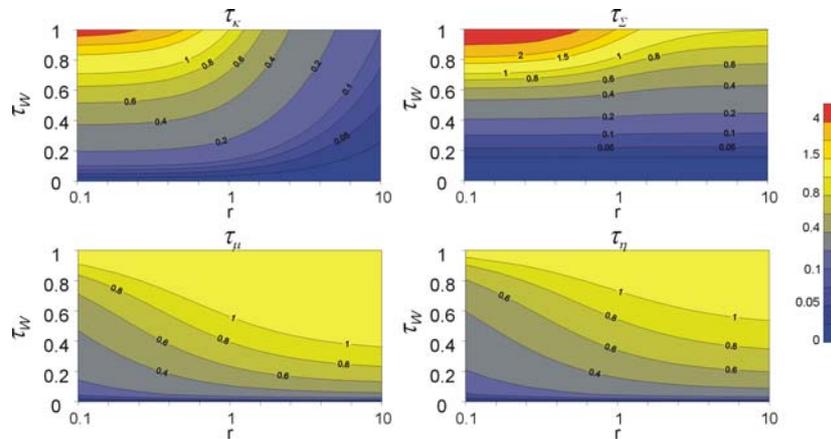
Figure 4.1: Bahr thresholds values obtained analytically for different values of the WAL threshold (τ_W) and r. Note that τ_W is represented on a linear scale and r on a logarithmic scale.

It can be observed that, for low values of r, τ_κ can be approximated to the constant value $\tau_\kappa \approx \dfrac{\tau_W}{\sqrt{2}}$, whereas for higher values of r its value decreases with r.

τ_μ and τ_η, which keep a $\sqrt{3/2}$ relation (eqs. 4.8b and 4.8c), smoothly increase in relation to both τ_W and r.

Threshold τ_Σ, except for high values of τ_W, has a low dependence with r, and take a constant value, $\tau_\Sigma \approx \tau_W^2$.

For a more accurate approach, instead of defining the new thresholds analytically, a statistical analysis was performed.

b) Statistical Development

For each dimensionality case, random values of the invariants, above or below their corresponding threshold range, τ_W, were generated, and Bahr parameters were computed using the analytical relationships in equations 4.2a to 4.2d. The value of r was chosen to be r=1, which is a valid approximation since, for experimental data, I_1 and I_2 are of the same order of magnitude (r represents the relation between the imaginary and real parts of the magnetotelluric tensor components). Thus, it was possible to establish the new thresholds for Bahr parameters, τ_κ, τ_μ, τ_η and τ_Σ, as well as the new dimensionality conditions. This development was made for different values of τ_W.

The following results were observed:

1) Thresholds τ_κ and τ_Σ can easily be related to τ_W using the approaches $\tau_\kappa \approx \dfrac{\tau_W}{\sqrt{2}}$ and $\tau_\Sigma \approx \tau_W^2$, as shown by the analytical development (eqs. 4.8a and 4.8d using r=1). The threshold τ_κ differentiates 1D and 2D from the rest of the cases, except 3D, which can take any value of κ. Values of Σ greater than τ_Σ indicate that the structure is not 1D, but 2D, with or without galvanic distortion (cases 2, 3 and 4). 3D cases can have any value of Σ.

2) The dependence of τ_μ on τ_W is more complex, since these parameters become more sensitive to the value of r, and, consequently the approximation r=1 is not the most appropriate.

In order to distinguish cases 1, 2 and 3 from case 4, there is a statistical value of τ_μ, although sometimes there is an overlap of possible values among these different cases.

3) In the case of parameter η, in order to differentiate between cases 4 and 5, it is also necessary to consider WAL parameter Q, which determines if I_7 is meaningless or not. This is accomplished assuming that the condition $\eta > \tau_\eta$ is only valid if Q is not meaningless, whereby

the threshold τ_Q is defined. The value of Q is also useful to differentiate case 3a from 3b, as in the WAL criteria.

Table 4.3 shows the values of τ_κ, τ_μ, τ_η and τ_Σ that correspond to r=1, τ_W= 0.1 and τ_Q=0.1. These thresholds, especially τ_Σ, differ significantly from the thresholds proposed by Bahr (1991).

Bahr parameters	τ_W=0.1
τ_κ	0.06
τ_μ	0.34
τ_η	0.12
τ_Σ	0.01

Table 4.3: Threshold values τ_κ, τ_μ, τ_η and τ_Σ, for r=1 , τ_W= 0.1 and τ_Q=0.1.

4.5 Bahr-Q Method

From the new threshold values obtained for Bahr parameters, the appropriate dimensionality conditions can be defined. However, there is a limitation by using only four parameters, compared to the five of the WAL method. In this section, the Bahr method is extended to be internally consistent with WAL invariants.

To classify dimensionality it is necessary to take into account parameters κ, μ, η, Σ and the thresholds as defined in the previous section. Also, in order to distinguish the two subcategories of case 3, it is necessary to use parameter Q. This new method has been named the Bahr-Q (B-Q) method. Table 4.4 shows the conditions necessary to classify the dimensionality using B-Q parameters (κ, μ, η, Σ and Q) with the thresholds values defined in the preceding section.

The application of the B-Q method in the example from Table 4.2 leads to the same interpretation as does WAL criteria: $\kappa = \mu = \eta = 0$, $\Sigma > \tau_\Sigma$ and Q>τ_Q, which indicates two-dimensionality.

B-Q method					
Dimensionality Case	κ	μ	Σ	η	Q
1) 1D	$<\tau_\kappa$	$<\tau_\mu$	$<\tau_\Sigma$	$<\tau_\eta$	---
2) 2D	$<\tau_\kappa$	$<\tau_\mu$	$>\tau_\Sigma$	$>\tau_\eta$ $<\tau_\eta$ $>\tau_\eta$	Q<τ_Q --- Q<τ_Q
3) a) 3D/2D twist b) 3D/1D2D	$>\tau_\kappa$	$<\tau_\mu$	$>\tau_\Sigma$	$<\tau_\eta$ ---	a) Q>τ_Q b) Q<τ_Q
4) 3D/2D	$>\tau_\kappa$	$>\tau_\mu$	$>\tau_\Sigma$	$<\tau_\eta$ $>\tau_\eta$	--- Q<τ_Q
5) 3D	---	---	---	$>\tau_\eta$	Q>τ_Q

Table 4.4: The Bahr-Q method criteria to characterise geoelectric dimensionality.

4.6 Application to Synthetic and Real Data

Three sets of data were used in order to compare the dimensionality obtained using WAL, Bahr and B-Q methods. The first is a synthetic set from the model used in Weaver *et al.* (2000). The second and third sets, increasing in complexity, consist of real data from the COPROD dataset, already used in the previous chapter, and from the BC87 dataset, also well known by the MT community. In the BC87 datasets, errors were taken into account.

Set 1: Synthetic Data

Set 1 consists of seven magnetotelluric tensors (A-G) selected from the synthetic model used in Weaver *et al.* (2000). It consists of a cubic conductive anomaly (0.5 $\Omega \cdot$m) embedded on the surface of an otherwise 2D structure, formed by a vertical fault that separates a layered medium of 10 $\Omega \cdot$m, 100 $\Omega \cdot$m and 1 $\Omega \cdot$m from a homogeneous medium of 1$\Omega \cdot$m (Figure 4.2).
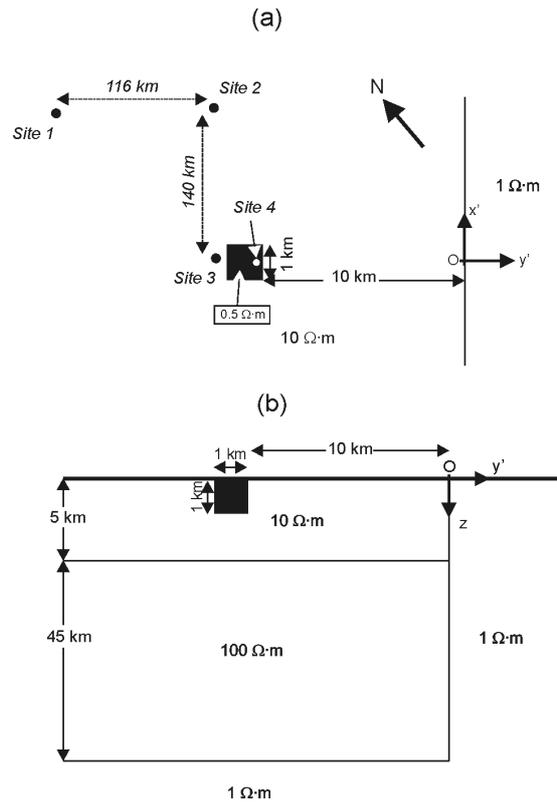
Figure 4.2: Synthetic model from Weaver *et al.* (2000). a) Plane view at z=0, b) Vertical cross section at x'=0.

The magnetotelluric tensors computed from the model for four sites at different periods exemplify the different types of dimensionality that can be identified using the WAL criteria with a threshold value of 0.1:

A: Site 1, 100s: 1D;

B: Site 2, 1000s: 2D with equal phases for xy and yx;

C: Site 2, 100s: 2D with different phases for xy and yx;

D: Site 3, 1000s; 3D/1D2D;

E: Site 2, 100s where E field was rotated 10°: 3D/2Dtwist;

F: Site 3, 100s: 3D/2D;

G: Site 4, 1s: 3D.

Table 4.5 presents the values of Bahr parameters and Q for each tensor, showing which are above (grey background) or below the threshold.

MT Tensor	κ (τ_κ=0.06)	μ (τ_μ=0.34)	Σ (τ_Σ=0.01)	η (τ_η=0.12)	Q (τ_Q=0.1)
A (1D)	0	0	0	0	0
B (2D)	0	0	0.09	0	0.01
C (2D)	0	0	0.05	0	0.36
D (3D/1D2D)	0.13	0.07	0.25	0.01	0.03
E (3D/2D twist)	0.18	0.02	0.05	0.01	0.36
F (3D/2D)	0.09	0.37	0.20	0.06	0.31
G (3D)	0.13	0.25	0.21	0.17	0.28

Table 4.5: Bahr parameters and Q, with their corresponding thresholds values, derived from τ_W=0.1, for tensors A-G from Set 1 (in parenthesis, the dimensionality according to the WAL method). In grey, parameters values greater than the threshold.

Table 4.6 shows the dimensionality that one would obtain using two types of analysis:

1) The Bahr method

2) The B-Q method: Bahr parameters and invariant Q, using τ_κ, τ_μ, τ_η and τ_Σ from τ_W=0.1 (see Table 4.3). Q is considered meaningless if its value is below 0.1.

The results present some equivalences but also some divergences showing how the B-Q method proposed improves the classical Bahr method. They also prove the consistency between B-Q and WAL methods.

MT Tensor (WAL dimensionality)	Bahr method	B-Q method
A (1D)	1D	1D
B (2D)	1D	2D
C (2D)	1D	2D
D (3D/1D2D)	3D/2D	3D/1D2D
E (3D/2D twist)	3D/1D	3D/2D twist
F (3D/2D)	2D	3D/2D
G (3D)	?	3D

Table 4.6: Dimensionality obtained for tensors A-G using the Bahr method and B-Q method.

In A (1D according to WAL) both criteria agree because the Bahr parameters are below the established thresholds, which indicate that the structure is 1D.

B and C data (2D according to WAL) are one-dimensional according to Bahr, whilst the magnetotelluric tensor has non-zero diagonal components and a well-defined strike direction (Weaver *et al.*, 2000). Using the B-Q method, these tensors are interpreted as 2D. This inconsistency between Bahr and B-Q methods is due to the new definition of τ_Σ, which is the square of τ_W. However, Bahr's method uses τ_Σ=0.1, which would correspond to a higher value for τ_W (0.31). As a result, tensors that are clearly 2D, may be interpreted as 1D.

In D (3D/1D2D according to WAL), the Bahr method would infer the following:

1) The structure is not 1D ($\kappa > 0.1$);

2) There is a phase difference between the tensor components ($\mu > 0.05$);

3) It can be interpreted as a distortion over a 1D or 2D structure ($\eta < 0.05$) and

4) The regional structure is 2D ($\Sigma > 0.1$).

According to B-Q, the distortion occurs over a 1D or 2D structure with equal phases in the non diagonal components of the MT tensor. Now, the inconsistency between Bahr and B-Q comes from the value of τ_μ: the value 0.05 established by Bahr (1991) is too low, and hence provides a very strict criterion to consider when phase values are different (less than 3° in this case).

In E (3D/2Dtwist), with the exception of κ, all the parameters are below the thresholds defined by Bahr and the structure could be considered 3D/1D. With the B-Q method, since Q is not meaningless, it corresponds to a 3D/2D structure, affected only by twist.

It must be noted that the use of parameter Q in the B-Q method allows one to distinguish two types of dimensionalities, 3D/1D2D and 3D/2Dtwist, in D and E.

In F (3D/2D according to WAL), the Bahr method states that the dimensionality is 2D, because both κ and Σ are greater than 0.1. B-Q agrees with a 3D/2D interpretation because only η is below the threshold.

In G (3D according to WAL), the Bahr method cannot ascertain if the structure is 3D because η is not greater than 0.3. By contrast, the B-Q method identifies the structure as 3D because $\eta > \tau_\eta$ and Q > 0.1.

From this analysis it can be concluded that the B-Q method, consisting of the use of Bahr parameters and WAL invariant Q, with the new conditions and a suitable threshold for each invariant, is consistent with WAL invariants criteria, used to determine the dimensionality of the magnetotelluric tensor.

Set 2: COPROD2 Dataset: Site 85_314

As it was shown in the previous chapter, the data responses of site 85_314 from the COPROD2 dataset (Appendix B) have in general a 1D behaviour at short periods (up to 10s). From 10s to 1000s the data display a 2D behaviour with a NS strike direction (Jones *et al.*, 2005), as determined from the multisite-multifrequency decomposition code of McNeice and Jones (2001), based on Groom and Bailey (1989) (G&B).

Table 4.7 shows Bahr and Q parameters related to the threshold values and the dimensionality derived from the B-Q method, compared to the ones obtained using G&B decomposition, WAL and Bahr methods.

The Bahr and G&B methods provide the simplest dimensionality descriptions, whereas WAL and B-Q also depict 3D cases at the longest periods. The WAL and B-Q results are coincident, describing the dimensionality as 1D up to 20s, 2D from 20s to 300s and 3D up to 1000s. G&B decomposition infers an optimum 2D description from 10 to 1000s and the Bahr method results in a 1D description up to 100s and 2D for the remainder.

Table 4.7: Dimensionalities obtained from the Groom and Bailey, Bahr, WAL and B-Q methods for site 85_314, COPROD2 set. The B-Q method uses τ_κ, τ_μ, τ_η and τ_Σ obtained from τ_{WAL}=0.1, and τ_Q=0.1 (Table 4.3). The grey area represents the range of periods for which G&B decomposition resulted in large errors.

Hence, all methods coincide with a 1D description up to 10s. G&B decomposition at this period range provided large error values, which is consistent with 1D dimensionality. The main difference is in the period at which the transition between 1D and 2D takes place, which is significant (about 1 decade with respect to the other methods) for the Bahr method. Furthermore, the 3D dimensionality cases at the longest periods are only shown in the WAL and B-Q methods.

The lowest misfit on the 3D/2D decompositions (Strike program) are obtained when the range 10s-300s is considered (B-Q method), as opposed to the 10s-1000s (Jones *et al.*, 2005) and 100s-1000s (Bahr Method). These results, together with large values of the invariant I_7 observed at the longest periods, confirm the validity of the dimensionality description obtained through the B-Q method.

Set 3: BC87 Dataset: Site 4

The BC87 dataset was acquired in British Columbia (Appendix C) and is commonly used to test and compare new methods in analysis and interpretation of MT data. The data display complex 3D effects, due both to local effects and the presence of the Nelson batholithic body (western part of profile). Site 4 is located above this body.

Table 4.8 summarises the kind of structures derived from the new B-Q method and compares them to those obtained using the WAL and Bahr methods, in both cases departing from τ_W=0.1, τ_Q=0.1 and without considering data errors. The WAL and B-Q methods give the same dimensionality interpretation, except for some periods.

The differences between the B-Q and WAL methods are due to parameters η and Q that have values close to the thresholds, which, as stated before, were not defined analytically but, rather, statistically.

B-Q and WAL interpretation is more complete than that of using the classical Bahr method, whose conditions, $\eta < 0.05$ and $\eta > 0.3$, do not allow the identification of data affected by distortion and 3D cases.

The dimensionality obtained using WAL and B-Q shows that up to 1s distortion can be removed from the data, which can be interpreted as 2D. For periods longer than 1s the dimensionality is 3D, with the exception of some particular periods, where 3D/2D cases are obtained.

This description agrees, in general, with Jones *et al.* (1993). Accordingly, a 3D/2D decomposition of the data is possible for periods shorter than 1s and longer than 10s. From 1s to 10s the data show a 3D behaviour, due to the presence of the Nelson Batholith. The difference with B-Q and WAL dimensionality results becomes evident at longer periods, where these

methods identify the data as 3D, whereas Jones *et al.* (1993) make a 3D/2D decomposition, with moderate misfit values in the retrieval of the regional responses.

κ and Σ (τ_κ=0.06) (τ_Σ=0.01)	μ (τ_μ=0.34)	η (τ_η=0.12)	Q (τ_Q=0.1)	Bahr Method	B-Q Method	WAL Method
	$>\tau_\mu$	$>\tau_\eta$			3D	3D
		$<\tau_\eta$			3D/2Dtwist	3D/2D twist
		$>\tau_\eta$			3D	3D
	$<\tau_\mu$	$<\tau_\eta$	$>\tau_Q$	3D/2D	3D/2Dtwist	3D/2Dtwist
$>\tau_\kappa$						3D
		$>\tau_\eta$			3D	
		$<\tau_\eta$	$<\tau_Q$		3D/1D2D	3D/2Dtwist
	$>\tau_\mu$				3D/2D	
	$<\tau_\mu$	$>\tau_\eta$			3D	3D
		$<\tau_\eta$			3D/2D	3D/2D
$>\tau_\Sigma$	$>\tau_\mu$	$>\tau_\eta$		3D	3D	3D
			$>\tau_Q$			
		$<\tau_\eta$			3D/2D	3D/2D
		$>\tau_\eta$			3D	3D
				3D		
				3D		
		$<\tau_\eta$			3D/2D	3D/2D

Table 4.8: B-Q parameters and dimensionality obtained from the Bahr, B-Q and WAL methods for the data from site 4, BC87 dataset. Parameters shaded in grey are those with values greater than the threshold.

When data errors are considered, which in turn influence the computed parameters, one of the consequences is that the dimensionality can be undetermined if the error bars cross the threshold values, since in these cases there are ambiguities in the classification. Another consequence is the bias that can appear between the true values (computed directly from the tensor components) and the statistical values (computed as the average of the different Gaussian generations) (chapter 3).

Figure 4.3 and Figure 4.4 show the values of the WAL and Bahr parameters and their errors that correspond to site 4 for two noise levels: 1% and 10%, where only half of the periods have been plotted. For WAL invariants I_7 and Q the biases are also presented (for the remainder of the parameters the biases are insignificant, as can be seen by comparing the statistical values for the different noise levels).
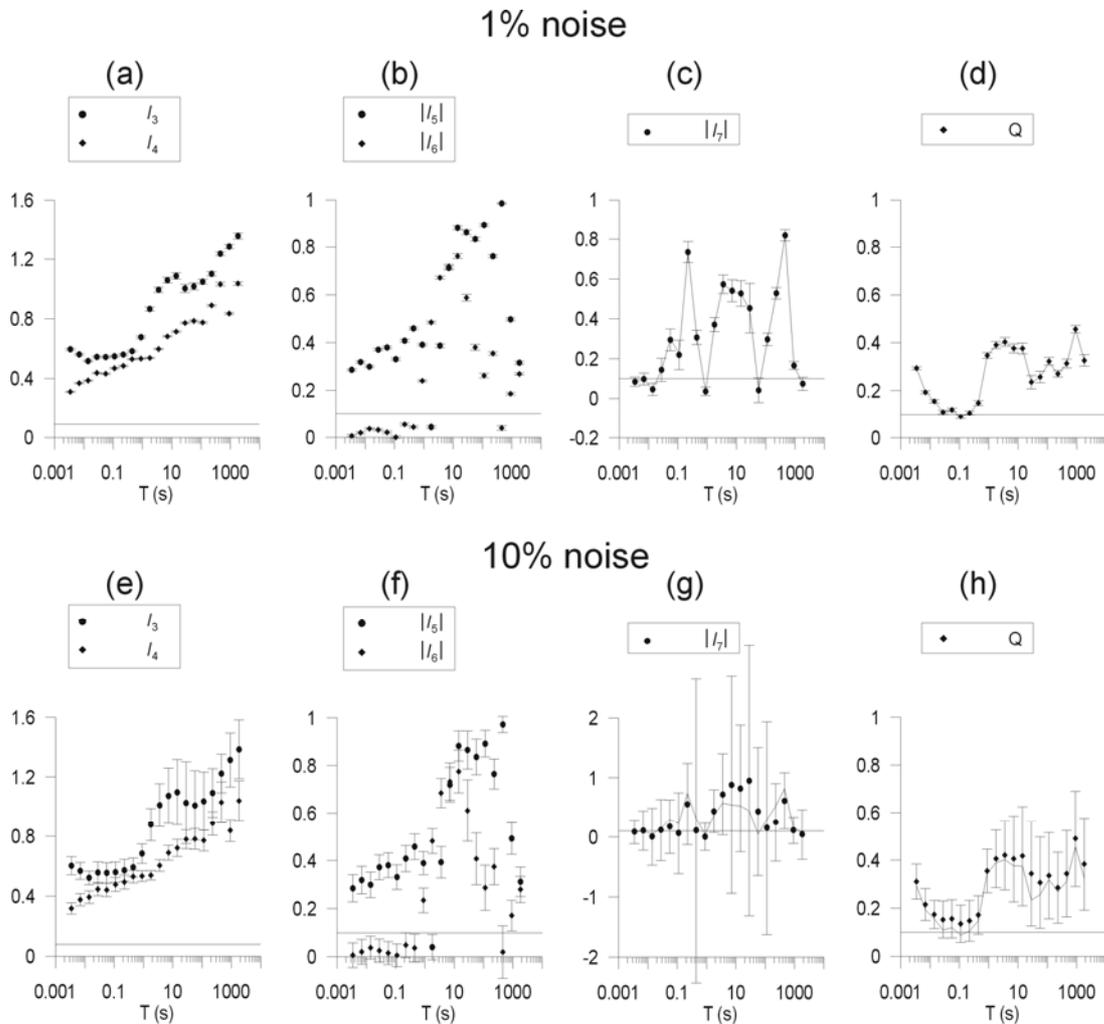
Figure 4.3: Invariants $I_3 - I_7$ and Q, with their error bars, for 20 recorded periods of site 4 from BC87 dataset. 1% noise level (a, b, c and d) and 10% noise level (e, f, g and h). For invariants I_7 and Q (d and h) the biases between the true and statistical values are also presented (continuous line). Horizontal continuous line: threshold value $\tau_W=0.1$.

The error bars of WAL parameters are, in general terms, proportional to the noise level, with the exception of I_7, which displays large error bars at some periods. With 10% noise, at the third last period of I_6, most of the periods of I_7 and some periods of Q, the error bars cross the threshold value, with the consequence that it is not possible to discern whether these invariants can be considered null or not. A second problem is caused by the biases of I_7, which add more uncertainty in the parameter estimation. On the contrary, the biases of Q are not important.

The error bars of Bahr parameters are smaller and, apart from the lowest periods of η with 10% noise, it is always possible to discern if Bahr parameters are above or below the threshold.
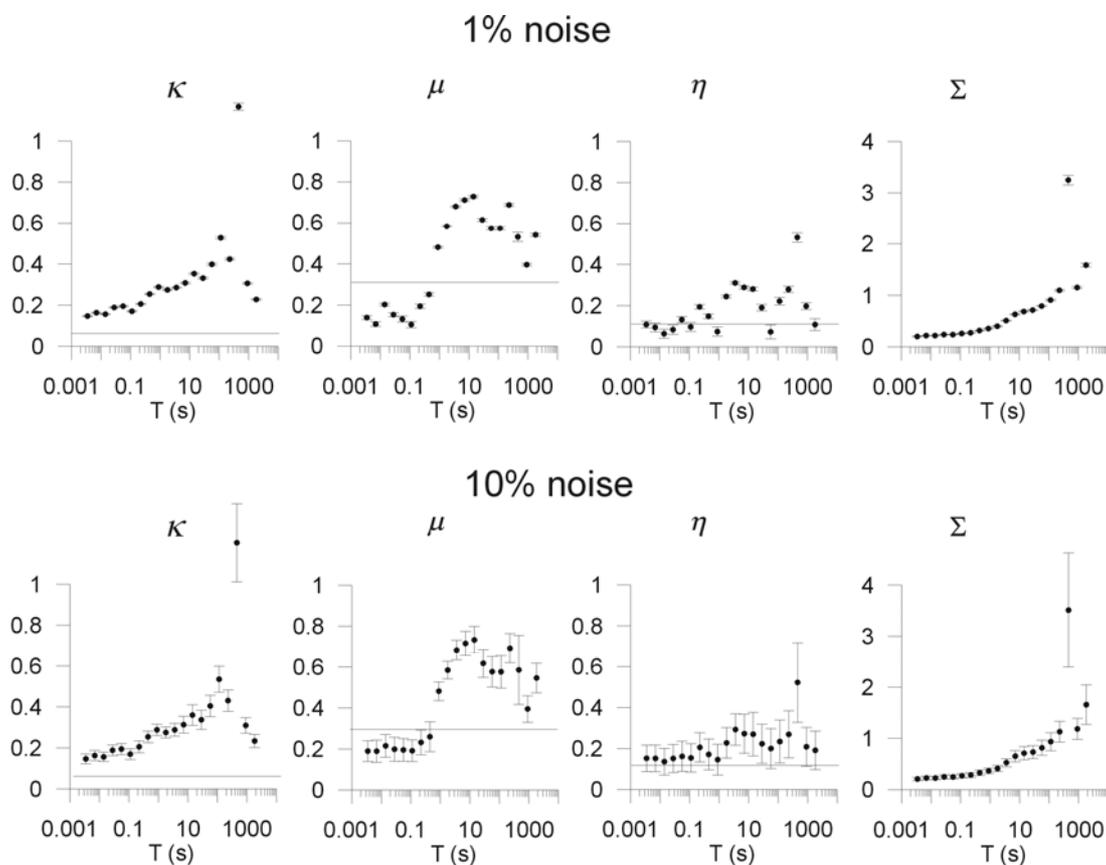
Figure 4.4: Bahr parameters with their error bars, for 20 recorded periods of site 4 from the BC87 dataset. Upper line: 1% noise. Lower line: 10% noise.

Consequently, having contrasted the large errors and biases of invariant I_7 versus the stability of Bahr and Q parameters, it can be stated that the Bahr-Q method is more robust under the presence of errors, in determining dimensionality.

A complete analysis of the dimensionality using the WAL and B-Q methods taking into account these errors would lead to a good determination of all periods for 1% noise. When the noise level is 10%, WAL invariant I_7 is undetermined at most periods. It limits the determination of the type of dimensionality using WAL criteria to only a few periods. With the same 10% noise level the Bahr-Q method provides a dimensionality description for a broader range of periods which, as already observed in the error bars, makes it more suitable for dimensionality analysis of noisy data.

4.7 Conclusions

In this chapter, two of the methods used to characterise the geoelectric dimensionality from MT data (the Bahr method and the WAL method) were studied. An important conclusion is that, compared to WAL, the Bahr method does not provide complete dimensionality interpretations. The reasons are that it uses only four parameters (whilst WAL uses five) and that the thresholds' values suggested lead to ambiguities or to situations where the dimensionality cannot be determined.

The redefinition of Bahr thresholds and the addition of the invariant Q allowed the establishment of a new dimensionality criterion (Bahr-Q) consistent with the WAL method.

The Bahr-Q method was tested with synthetic and real data corresponding to all types of dimensionality, obtaining the same interpretation as the WAL method. When the data are affected by a significant level of noise (10% or higher), the errors and biases of Q can be important and, consequently, the dimensionality obtained from either method can be undetermined. However, Bahr and Q parameters are more stable under noise effects and make Bahr-Q a more robust method.

Chapter 5: Applications of the Magnetotelluric Phase Tensor and Comparison with other Methods

The goal of this chapter is to analyse the phase tensor proposed by Caldwell *et al.* (2004) and apply it to synthetic and real data in order to determine the resolution of the parameters involved.

The phase tensor and its associated parameters and errors were analysed, and, together with the information obtained, compared to the WAL and Groom and Bailey methods. A last step consisted of fitting the phase tensor information to a 2D description of the data.

The phase tensor, and its related parameters and graphic representations were computed using a code created with Compaq Visual Fortran 2000® and a Matlab® application provided by Hugh Bibby.

5.1 The Phase Tensor for Different Dimensionality Cases

The phase tensor (see chapter 2, section 2.6) displays distinct expressions for particular cases of dimensionality and orientations of the measuring axes. The expressions of the phase tensors with their SVD (Singular Value Decomposition) and related parameters were computed for four particular cases (1D, 2D and 3D/2D with different orientations, and 3D), which are summarized, together with a numerical example in each case (Table 5.1). Note that the 2D example (B) is the same MT tensor used in chapter 4 (Table 4.2), which is rotated 30° in example C.

A	1D
Impedance tensor	$\begin{bmatrix} 0 & M \\ -M & 0 \end{bmatrix}$
Phase tensor	$\begin{bmatrix} -Y/-X & 0 \\ 0 & Y/X \end{bmatrix} = \begin{bmatrix} \tan(\varphi \pm \pi) & 0 \\ 0 & \tan\varphi \end{bmatrix}$
SVD Parameters	$\alpha_P = 0/0 \;\; (undefined), \Phi_{\substack{Max \\ min}} = \tan\varphi \;; \; \beta_P = 0^\circ$
Numerical example	**1D with φ=26.56º**
Impedance tensor	$\begin{bmatrix} 0 & 10+5i \\ -10-5i & 0 \end{bmatrix}$
Phase tensor	$\begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix} = \begin{bmatrix} \tan(26.56^o) & 0 \\ 0 & \tan(206.56^o) \end{bmatrix}$
SVD	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
SVD Parameters	$\alpha_P = 0/0 \;\; (undefined) ; \Phi_{\substack{Max \\ min}} = 0.5 \;; \; \beta_P = 0^\circ$

Table 5.1: Synopsis and numerical examples of the phase tensor and SVD parameters for particular dimensionality cases: (a) 1D case.

B	2D or 3D/2D along strike direction	
Impedance tensor	$\begin{bmatrix} 0 & M_{12} \\ M_{21} & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & M_{TE} \\ M_{TM} & 0 \end{bmatrix}$
Phase tensor	$\begin{bmatrix} Y_{21}/X_{21} & 0 \\ 0 & Y_{12}/X_{12} \end{bmatrix} = \begin{bmatrix} \tan\varphi_{21} & 0 \\ 0 & \tan\varphi_{12} \end{bmatrix}$	$\begin{bmatrix} Y_{TM}/X_{TM} & 0 \\ 0 & Y_{TE}/X_{TE} \end{bmatrix} = \begin{bmatrix} \tan\varphi_{TM} & 0 \\ 0 & \tan\varphi_{TE} \end{bmatrix}$
SVD Parameters	$\alpha_P = 0^\circ, \; \Phi_{\substack{Max \\ min}} = \tan\varphi_{\substack{12 \\ 21}}, \beta_P = 0^\circ$	
Numerical example	**2D along strike direction (example from chapter 4, table 4.2)**	
Impedance tensor	$\begin{bmatrix} 0 & 25+9i \\ -15-12i & 0 \end{bmatrix}$	
Phase tensor	$\begin{bmatrix} 0.8 & 0 \\ 0 & 0.36 \end{bmatrix} = \begin{bmatrix} \tan 38.66^o & 0 \\ 0 & \tan 19.79^o \end{bmatrix}$	
SVD	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0.8 & 0 \\ 0 & 0.36 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	
SVD Parameters	$\alpha_P = 0^\circ ; \; \begin{array}{l} \Phi_{Max} = 0.8 = \tan(38.66^o) \\ \Phi_{min} = 0.36 = \tan(19.79^o) \end{array} ; \; \beta_P = 0^\circ$	

Table 5.1 (cont.) (b) 2D or 3D/2D measured along the strike

C	2D or 3D/2D rotated an angle θ clockwise
Impedance tensor	$\begin{bmatrix} (M_{12}+M_{21})\sin\theta\cos\theta & M_{12}\cos^2\theta - M_{21}\sin^2\theta \\ -M_{12}\sin^2\theta + M_{21}\cos^2\theta & -(M_{12}+M_{21})\sin\theta\cos\theta \end{bmatrix}$
Phase tensor	$\begin{bmatrix} (Y_{21}/X_{21})\cos^2\theta + (Y_{12}/X_{12})\sin^2\theta & (-Y_{21}/X_{21}+Y_{12}/X_{12})\sin\theta\cos\theta \\ (-Y_{21}/X_{21}+Y_{12}/X_{12})\sin\theta\cos\theta & (Y_{21}/X_{21})\sin^2\theta + (Y_{12}/X_{12})\cos^2\theta \end{bmatrix} =$ $\begin{bmatrix} \Phi_{11}\cos^2\theta + \Phi_{22}\sin^2\theta & (-\Phi_{11}+\Phi_{22})\sin\theta\cos\theta \\ (-\Phi_{11}+\Phi_{22})\sin\theta\cos\theta & \Phi_{11}\sin^2\theta + \Phi_{22}\cos^2\theta \end{bmatrix}$
SVD Parameters	$\alpha_P = -\theta$, $\Phi_{\substack{Max \\ min}} = \tan\varphi_{\substack{12 \\ 21}}$, $\beta_P = 0^\circ$
Numerical example	**2D (example B) rotated 30º**
Impedance tensor	$\begin{bmatrix} 4.33-1.29i & 22.5+9.75i \\ -17.5-11.25i & -4.33+1.29i \end{bmatrix}$
Phase tensor	$\begin{bmatrix} 0.69 & -0.1905 \\ -0.1905 & 0.47 \end{bmatrix}$
SVD	$\begin{bmatrix} 0.866 & 0.5 \\ -0.5 & 0.866 \end{bmatrix}\begin{bmatrix} 0.8 & 0 \\ 0 & 0.36 \end{bmatrix}\begin{bmatrix} 0.866 & -0.5 \\ 0.5 & 0.866 \end{bmatrix}$
SVD Parameters	$\alpha_P = -30^\circ$; $\begin{array}{c}\Phi_{Max} = 0.8 = \tan(38.66^\circ)\\ \Phi_{min} = 0.36 = \tan(19.79^\circ)\end{array}$; $\beta_P = 0^\circ$

Table 5.1 (cont.) (c) 2D or 3D/2D rotated

D	3D case
Impedance tensor	$\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}$
Phase tensor	General expression (eq. 2.28)
SVD Parameters	General expressions (eqs. 2.31, 2.32 and 2.33)
Numerical example	**3D case**
Impedance Tensor	$\begin{bmatrix} 1.405+2.23i & 5.33+2.5i \\ -7.45-4.23i & 1.45+3.29i \end{bmatrix}$
Phase Tensor	$\begin{bmatrix} 0.617 & -0.333 \\ 0.256 & 0.557 \end{bmatrix}$
SVD	$\begin{bmatrix} 0.97 & 0.22 \\ -0.22 & 0.97 \end{bmatrix}\begin{bmatrix} 0.7061 & 0 \\ 0 & 0.6076 \end{bmatrix}\begin{bmatrix} 0.77 & -0.6335 \\ 0.6335 & 0.77 \end{bmatrix}$
SVD parameters	$\alpha_P = -26^\circ$; $\begin{array}{c}\Phi_{Max} = 0.706 = \tan(35.22^\circ)\\ \Phi_{min} = 0.607 = \tan(31.28^\circ)\end{array}$; $\beta_P = -13^\circ$

Table 5.1 (cont.) (d) 3D case.

The phase tensor ellipses corresponding to the numerical examples A, B, C and D are represented in Figure 5.1. As expected, the 1D case (A) representation is a circle, with an undefined value of α_P. 2D cases (B and C) are ellipses with clearly differentiated major and minor axes, aligned along the strike direction. The 3D case (D) is characterized by almost equal values of major and minor ellipse axes, although with well defined values of α_P and β_P, which in this case both have a non-zero value.
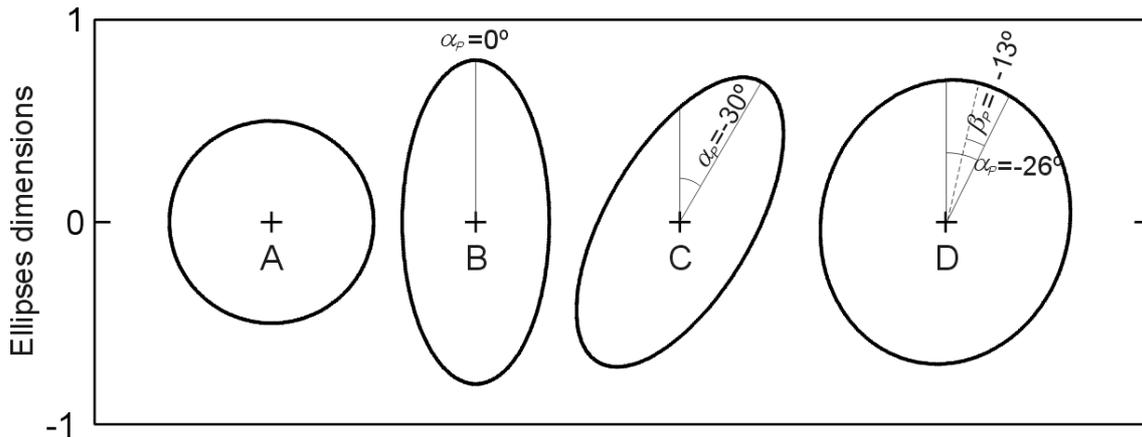
Figure 5.1: Representation of the phase tensor ellipses corresponding to the synthetic examples A, B, C and D displayed in Table 5.1.

5.2 Application to Synthetic and Real Datasets

5.2.1 Synthetic model dataset

The first dataset consists of the responses of a synthetic model, previously used in the work of Ledo *et al.* (2002), which studied the 3D effects in the 2D interpretation of magnetotelluric data. It consists of a 3D body embedded in a 2D structure (Figure 5.2).

The responses correspond to 30 sites along a profile, with 11 periods from 0.01 s to 1000 s. Random galvanic distortion C, with gain $g = 1$ and without anisotropy (equation 1.26), was added to each site in order to see the effects of distortion over 3D regional responses.

To make the responses representative of real data, Gaussian noise was added, which is not proportional to the signal amplitude. Thus, the same error, the relative error of the largest of the four components, was added to all tensor components. Additionally, the values of the components were randomly scattered around their errors, such that the final values of the tensor components were the means of these scattered values.

Once distortion was added to the responses, two datasets were created, each one with a different level of noise: dataset SIT-1, with 1% error, and SIT-2.5, with 2.5% error. The original dataset, without distortion or noise, was also employed and referred to as SIT.
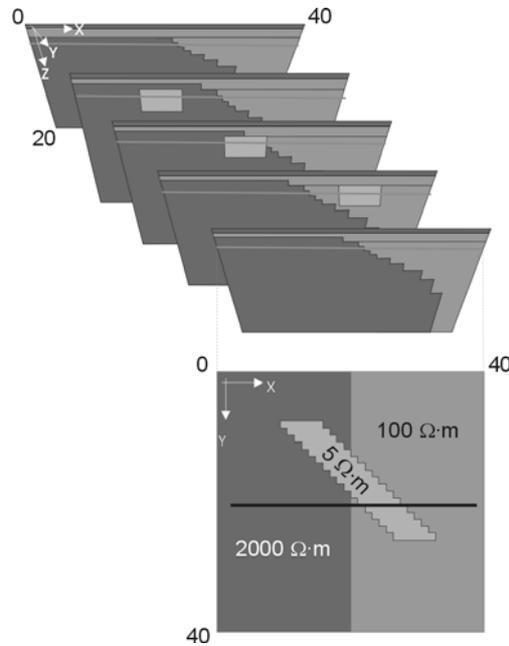
Figure 5.2: 3D electrical conductivity regional model used to generate synthetic responses. Black line on XY view indicates the position of the profile.

The estimations of the phase tensor parameters and their errors were computed using Random Gaussian Noise generation (statistical values and errors) (chapter 3, approach **b**), using n=100. The true values of the parameters were also considered.

The phase tensor ellipses of dataset SIT are displayed in Figure 5.3, together with the arrows indicating the direction given by angle α_P. Figure 5.4 represents the associated parameters Φ_{Max}, Φ_{min}, α_P, β_P, the difference Φ_{Max} - Φ_{min} and, except for the original dataset SIT, the error of angle α_P.

At the shortest and longest periods, the phase tensor representations (Figure 5.3) are circles, with different radii values, signifying changes in the phase values. These circles are characterised by low values of Φ_{Max} - Φ_{min} (Figure 5.4), and a wide variation of angle α_P along the different phase tensor representations. The angle β_P allows discerning whether the circles indicate a 1D or 3D structure. This angle has a range of values between -2° and 2°, which suggests that the circles can be interpreted as 1D.

At middle periods, the phase tensors are represented by general ellipses, with a gradual change in orientation, from the left side (90°, perpendicular to the model strike), to the right side (ellipses aligned along the model strike) of the profile. This change reflects the variations in the TE and TM phase values. In the same way as the other methods to determine the strike direction, the phase tensor representation also involves a 90° ambiguity in its determination.
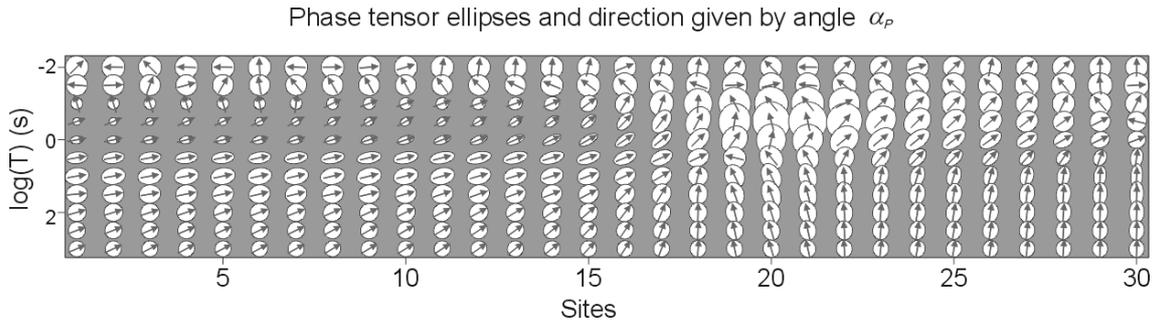
Figure 5.3: Ellipse representations of the phase tensor for SIT dataset responses, along with the direction given by α_P, plotted along the profile of the synthetic model for the 11 computed periods.

As for the datasets in which distortion and errors were added, at site 12 from SIT1 and sites 12, 20 and 25 from SIT2.5 there are important overflows of Φ_{Max} and, for all the parameters, the differences with the original dataset SIT are evident (Figure 5.4). These significant variations with respect to the original dataset rely on the high dependence of the phases on small changes in the values of the impedance tensor components, especially when their values are small. At the rest of sites, there is good agreement between the original and distorted data, as is expected from the property of invariance under distortion of the phase tensor. However, some remarkable points arise:

1) The statistical values of Φ_{Max} are in general higher than its true values, whereas those of Φ_{min} are smaller that the true ones. This implies that the difference between Φ_{Max} and Φ_{min} computed from the statistical values is greater than the true difference, which in some cases leads to a 2D misinterpretation of the data. This larger difference is a consequence of the errors, that results in a greater value of the radicand in eq. 2.33, consequential upon a larger value of the sum, Φ_{Max}, and a smaller value of the difference, Φ_{min}.
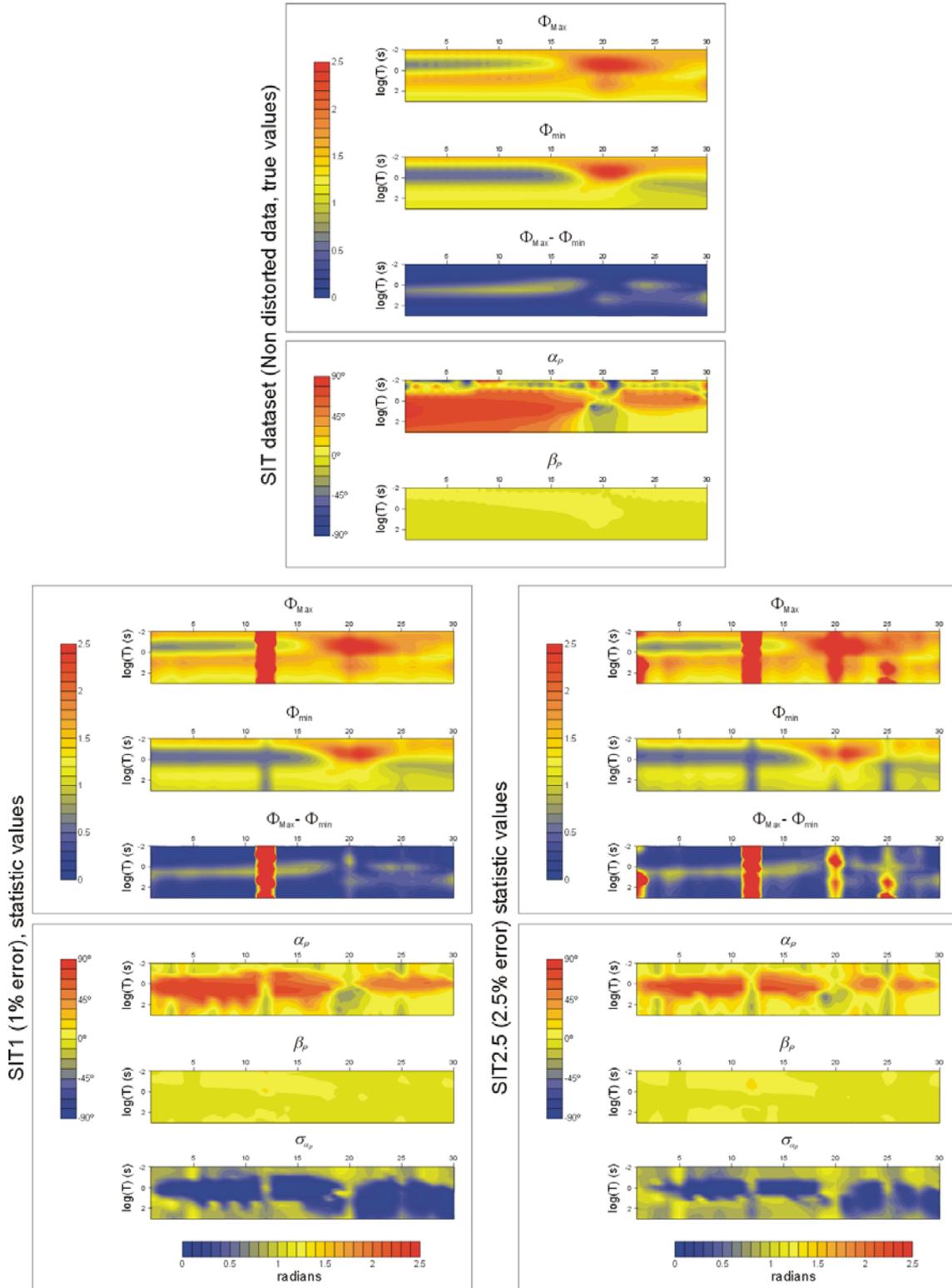
Figure 5.4: Cross sections of the phase tensor parameters corresponding to the three datasets computed from the synthetic model, plotted along the profile for the 11 computed periods. Upper panel: original dataset SIT, without errors or distortion. Lower panels: datasets in which distortion has been applied, and error added. Left: SIT1, with 1% error. Right: SIT2.5, with 2.5% error. The plotted parameters are Φ_{Max}, Φ_{min}, α_P, β_P, Φ_{Max}-Φ_{min} and σ_{α_P} (only for the datasets with errors).

2) The values of angle α_P for SIT1 and SIT2.5, at sites 1 to 15, present significant variations with respect to SIT, especially at long periods (Figure 5.4). Namely, whithout distortion, $\alpha_P \approx 90^o$, whereas for the distorted data $\alpha_P \approx 0^o$- 45°. This difference is explained by incorrect determination of this angle when the difference between Φ_{Max} and Φ_{min} is small. At these sites and periods, σ_{α_P} is greater than Φ_{Max}-Φ_{min}, so a 1D or 3D description of the data is more suitable. At middle periods, σ_{α_P} is small, so α_P is well determined and takes the same values for the original responses as for the two noise levels in the distorted data. At the shortest periods α_P is poorly resolved too, as is indicated by the small difference between Φ_{Max} and Φ_{min}.

As an illustration of how the bias between the statistical and true values affects the determination of angle α_P, Figure 5.5 represents the normalized biases of α_P ($bias_{norm} = \left| (\alpha_{Pstat} - \alpha_{Ptrue})/\alpha_{Ptrue} \right|$) for 1% and 2.5% noise. The areas with the greatest biases correspond to the areas with the largest error values, and these biases are higher for the SIT2.5 dataset. In both datasets, for about 50% of the data the normalised bias is greater than 1. The large bias at site 12 reflects the incorrect determination of α_P when distortion is added, since the regional data does have a well-determined α_P, which is constant at all periods.
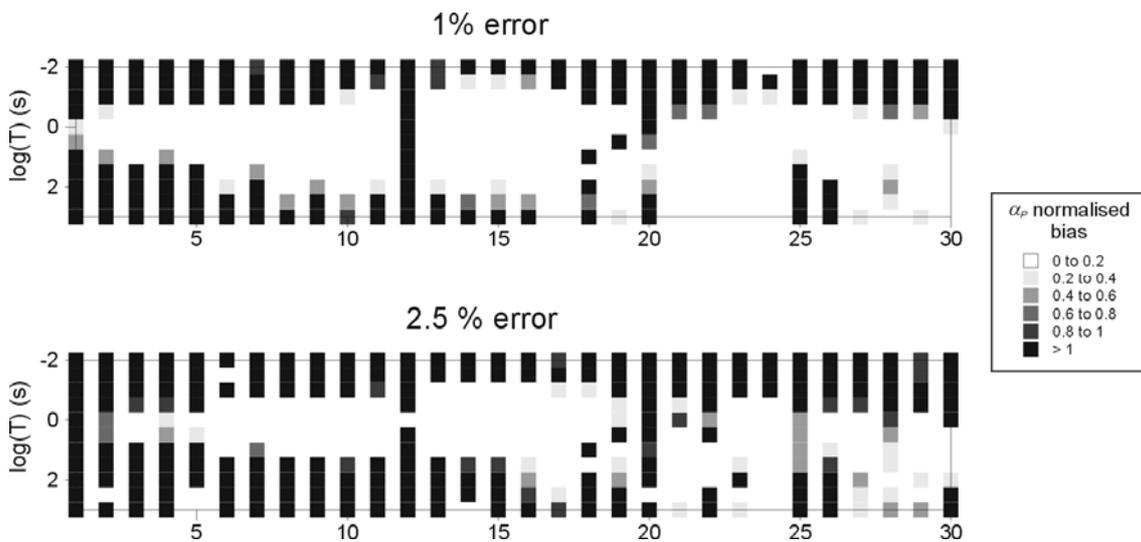
Figure 5.5: Normalised biases of α_P for SIT1 and SIT2.5 datasets. The normalised biased is computed as the difference between the statistical and true values divided by the true value of α_P.

Angle β_P presents values between -4° and 4° for all sites and periods. Both the statistical and true values are similar, and biases between true and statistical values are not significant. The central periods, in the left side of the profile, where β_P has values larger than 3° can be identified as 3D.

The overflows observed at sites 12, 20 and 25 were presumed to be caused by the galvanic distortion added at these sites, given that the noise level is the same for the whole dataset. These sites have shear angle values (φ_s) of 44°, -41° and 49° respectively, the closest to 45° or -45°, compared to the rest of sites.

Figure 5.6 represents, for 1% noise, parameters Φ_{Max}, Φ_{min}, α_P and β_P for each of the 100 realizations generated at T=1 second at sites 12 (φ_s = -44°) and 14 (φ_s = -5°), which have a similar geoelectric structure below but different distortion characteristics. Since the computed parameters should not be affected by distortion, and given the proximity between both sites and the same noise level, one would expect a similar behaviour for these parameters, but site 14 is highly stable whilst 12 is not. This observation reinforces the fact that, under extreme distortion conditions, in addition to noise effects, the phase tensor parameters become unstable and the distortion-invariance property is not valid.
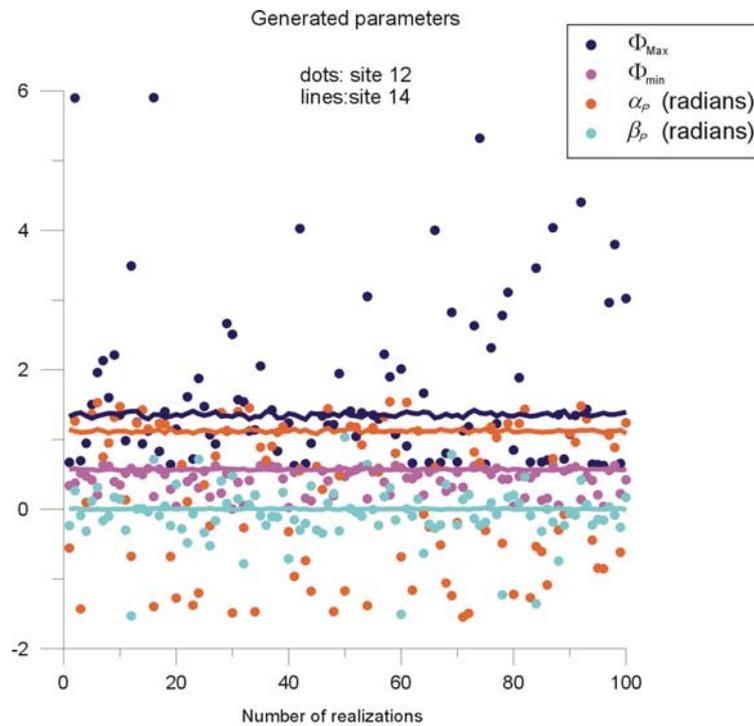
Figure 5.6: Values of phase tensor parameters for the 100 realizations obtained at site 12 and site 14 at T=1 s, with 1% error.

This first dataset allowed the characterization of the phase tensor under different conditions of distortion and noise. The parameters are highly sensitive to the errors and also to the type of dimensionality (e.g. α_P is highly scattered in 1D cases), and the effects of noise and distortion in the data lead to important differences with respect to the regional data.

5.2.2 The COPROD2 dataset: site 85_314

Figure 5.7 displays phase tensor representations of site 85_314 along the registered periods. It follows a simple pattern: circles for periods shorter than 10s, and ellipses with an approximated 60° inclination for the long periods (excluding the longest one). The values of β_P are small ($<3^o$) for the whole range of periods, with the exception of the longest one. Hence, it agrees with the description of the data representing a 1D dimensionality at short periods, 2D at the long ones with a 60° strike and 3D at the longest.
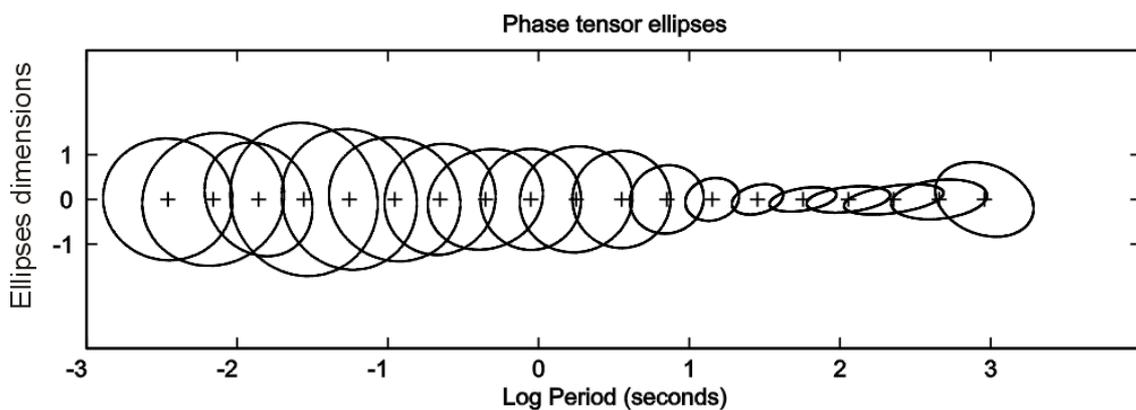
Figure 5.7: Phase tensor ellipses plotted for site 85_314 from COPROD2 dataset, for the different periods registered.

5.2.3 The BC87 dataset

The data corresponding to the BC87 profile was used to compute and analyse the phase tensor general features. Site 4 was also taken to study it in more detail.

Phase tensor parameters and errors were estimated using Random Gaussian Noise generation (as in the synthetic model dataset). No significant differences were found between the true and statistical values.

Phase tensor parameters Φ_{Max}, Φ_{min} were converted into their related phases, φ_{Max} and φ_{min} (equation 2.34). These phases are expressed in the first quadrant, so it must be taken into

account that one of these is in fact φ-180°, such as to agree with a 2D representation. Figure 5.8 presents the cross sections of φ_{Max} and φ_{min}, α_P and β_P. φ_{Max} and φ_{min} have values between 45° and 90°, and their differences (φ_{Max} - φ_{min}) are appreciable, but not too large. On the contrary, the supposed strike direction, given by α_P, changes abruptly for differing periods and from site to site. The angle β_P presents low values, except for long periods in the eastern part of the profile and middle periods in the western part.

In general, the phase tensor related angles are consistent with a 2D behaviour of the data. However, the abrupt changes in angle α_P and the local increases of β_P make it difficult to find a suitable strike direction to make a 2D interpretation of the profile.
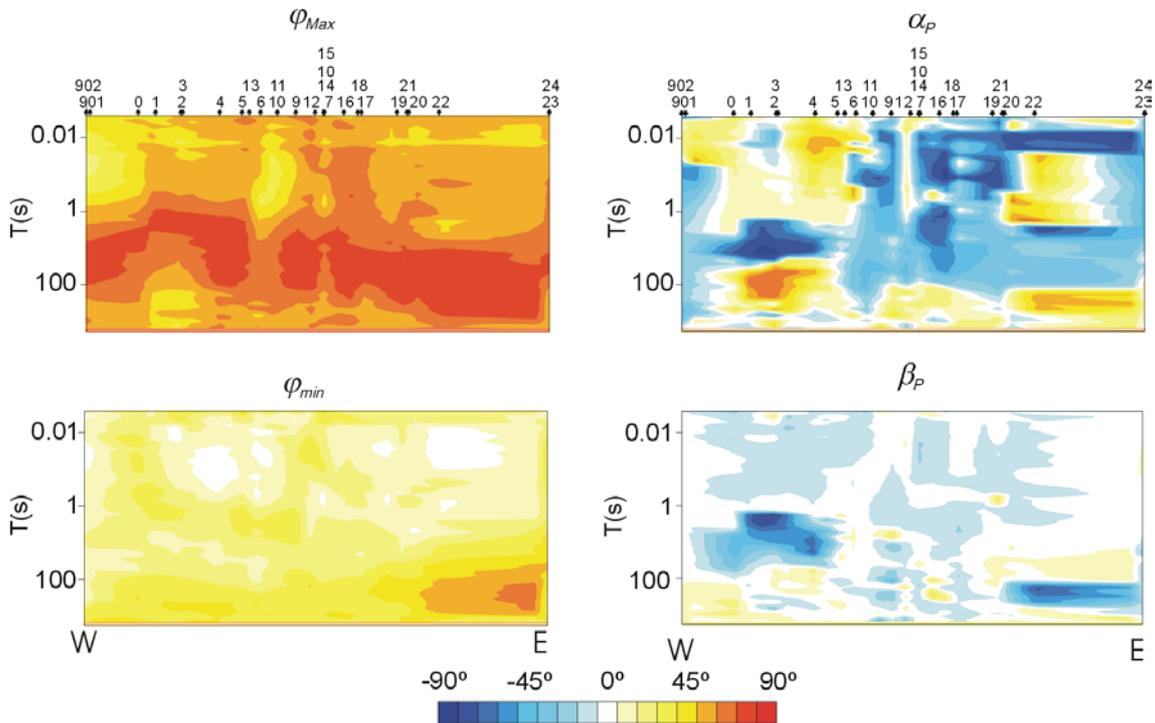
Figure 5.8: Cross sections of phase tensor related angles plotted along BC87 profile and the registered periods. φ_{Max} and φ_{min} are the arctangents of parameters Φ_{Max} and Φ_{min}.

Site 4, located over the Nelson Batholith, was studied in more detail. Figure 5.9 plots the phase tensor ellipses for the range of periods registered together with the direction given by angle α_P. If this direction differs from the maximum axes of the ellipse, a 2D interpretation is not possible, even if Φ_{Max} and Φ_{min} have different values.

Up to 2 s, the major axes of the ellipses coincide with α_P direction, which increases gradually from 45° to almost 90°. Hence, those periods in which the difference between Φ_{Max} and Φ_{min} is not meaningless can be interpreted as 2D (or 3D/2D).

From 2 s to 100 s, the major axes of the ellipses present large values (which mean large values of the phases, close to 90°), with clear divergences between the ellipse orientation and the direction given by α_P. At the longest periods the size of the ellipses decreases and Φ_{Max} and Φ_{min} are still clearly different but their orientation, however, still does not provide the direction of the strike. Consequently, for periods longer than 2s, the data must be treated, in general, as 3D.
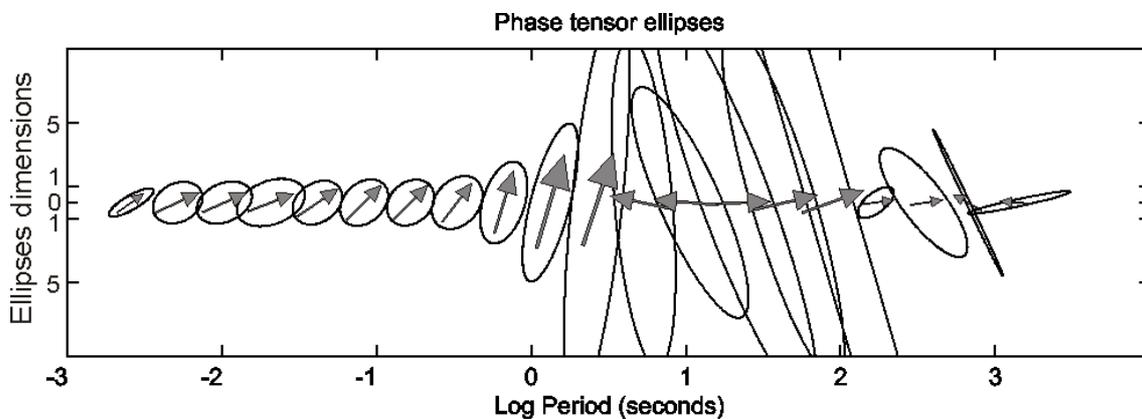
Figure 5.9: Phase tensor ellipses plotted for site 4 from the BC87 dataset for the different periods registered. Arrows indicate the direction given by angle α_P.

Figure 5.10 plots phase tensor angles φ_{Max}, φ_{min}, α_P and β_P with their error bars for site 4. The values of the parameters quantify and reinforce the description provided by the phase tensor representation.

The errors of φ_{Max} and φ_{min} are small, except for periods longer than 100s, although their error bars never cross, so they do not become significant. Errors in the direction angles, α_P and β_P, are small for the shortest periods, in which a 2D strike direction can be defined. These errors become important from T=2s onward, which agrees with a 3D description of the data, without any prevalent strike direction.
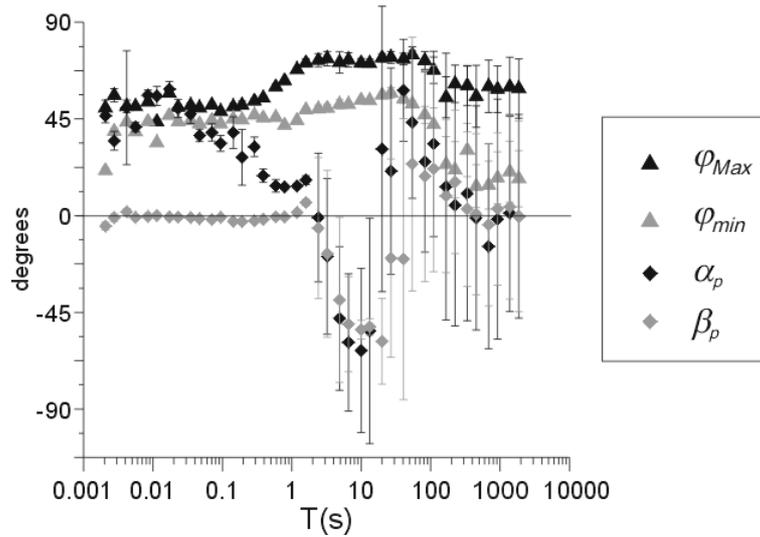
Figure 5.10: Plots of φ_{Max}, φ_{min}, α_P and β_P with their error bars for the registered periods of site 4 from the BC87 dataset.

5.3 Comparison with the WAL Method

The equivalences and differences between the characterisation of the dimensionality from WAL invariants and the phase tensor have already been seen from the theory (section 2.6 and Caldwell *et al.*, 2004) and the examples used.

This section presents the main results of the comparison between the phase tensor and WAL invariants, using the data from section 5.2. This comparison focused on finding new equivalences between both methods and on the determination of the corresponding parameters, under noise and distortion conditions. Another study on the comparison between WAL and the phase tensor can be found in Weaver *et al.* (2003).

For the datasets from the synthetic model, SIT, SIT1 and SIT2.5, WAL and phase tensor parameters and the strike directions obtained from each method were computed and compared.

The cross-sections of Φ_{Max} - Φ_{min} and invariant Q for the SIT dataset (Figure 5.11) show a certain degree of proportionality between both parameters, both dimensionless. A second comparison between φ_{Max} - φ_{min} (the difference between the angles related to Φ_{Max} and Φ_{min}) and invariant Q shows the same values for both parameters (Figure 5.12a), even though φ_{Max} - φ_{min} is an angular magnitude, expressed in radians.

This similarity disappears when distortion is added (Figure 5.12b), since φ_{Max} - φ_{min} is distortion invariant (with some exceptions due to the overflows caused by some distortion parameters), whereas Q is not invariant under distortion and its value changes.
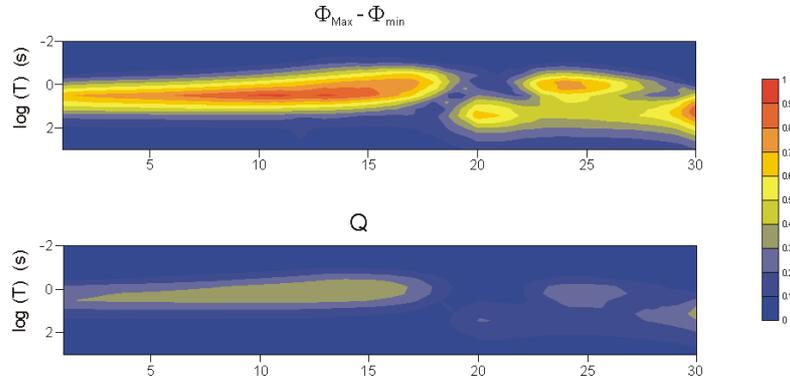
Figure 5.11: Cross-sections of Φ_{Max} - Φ_{min} (from the phase tensor) and WAL invariant Q, for the registered periods along the SIT dataset profile.
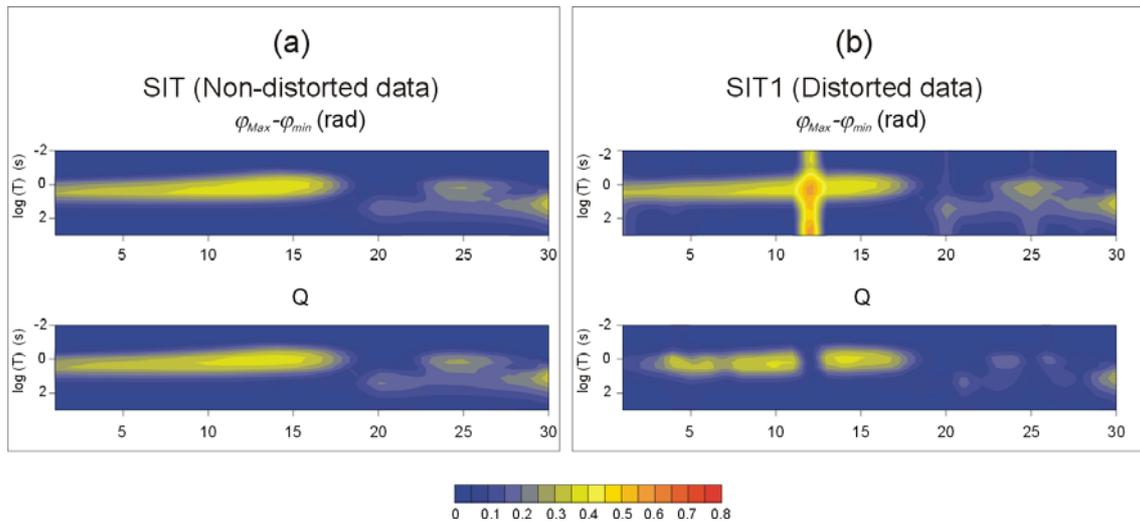
Figure 5.12: Cross-sections of φ_{Max} - φ_{min} (in radians, from the phase tensor) and WAL invariant Q, for the registered periods. a) SIT dataset (non-distorted data), b) SIT1 dataset (distorted data with 1% error added).

All the examples of section 5.1, to which distortion was added, confirmed this equivalence between Q and φ_{Max} - φ_{min} for non-distorted data.

For the real data from site 85_314 (COPROD2) and site 4 (BC87) the coincidence and non-coincidence between Q and φ_{Max} - φ_{min}, depending on data being affected of not by distortion, is less clear (Figure 5.13).

Up to 10 s, where according to the WAL and Bahr-Q methods the dimensionality is 1D for site 85_314 and 3D mixed with 3D/2D for site 4 (see Chapter 4, section 4.6), Q and φ_{Max} - φ_{min} have similar values, with the exception of some scattered points for site 85_314 and between 1s and 5s for site 4. For 10s and longer periods (2D at site 85_314 and again 3D mixed with 3D/2D at site 4), the difference between both parameters becomes larger, although at site 85_314 both follow the same trend, whereas at site 4 the differences are more irregular. In the first site, the differences may be attributed to noise effects, whereas in the second, these may be, in fact, due to the distortion.

This coincidence between Q and φ_{Max} - φ_{min} when data are not distorted is an important result, since it opens a new method for identifying galvanic distortion, even in 3D cases. Moreover, it has been tested successfully in synthetic data, where the distortion was important.
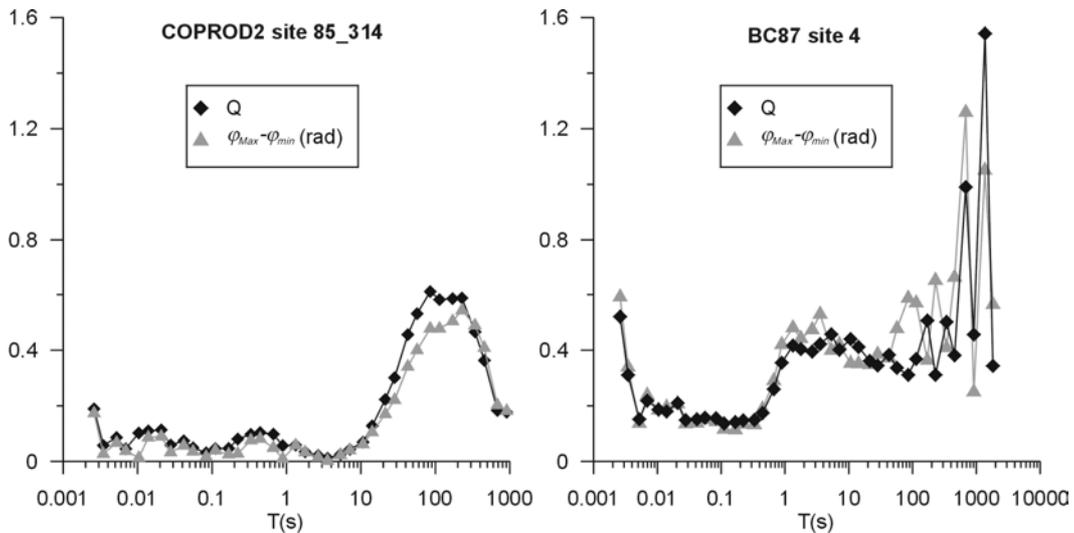
Figure 5.13: Graphical representations of Q invariant (from WAL) and φ_{Max} - φ_{min} for site 85_314 from the COPROD2 dataset and site 4 from the BC87.

In order to compare the determinations of the strike using both approaches, the strike angles were estimated for the SIT1 and SIT2.5 datasets as α_P -β_P from the phase tensor, and θ_3 (chapter 2, equation 2.23) from the WAL method. In both cases Random Gaussian Noise generation was applied. For each dataset, similar values of the strike directions were obtained. The main differences are in the errors, which are larger in the strike directions determined from

WAL parameters. This is a consequence of the different expressions of the strike angles, which leads to different error estimations.

5.4 Comparison with Groom and Bailey Decomposition

A comparison between the phase tensor and Groom and Bailey (G&B decomposition) (Groom and Bailey, 1989) was performed, using site 4 from BC87. Two G&B decompositions of this site are available from Jones *et al.*, 1993.

The first analysis, a) was performed assuming a 2D dimensionality, without distortion. The strike angle obtained is 45° up to 1 s and decreases towards 0° at the longest periods. φ'_{yx} reaches values greater than 90° at the longest periods, and the errors are large (Figure 5.14a).

Analysis b) assumed a 3D/2D dimensionality, with a 60° strike. Both φ'_{xy} and φ'_{yx} phases can be represented in the first quadrant, although a large difference between them is observed at periods from 1s to 10s. Distortion parameters show two different behaviours: twist and shear close to zero from 0.01 to 1s, and increasing to $\theta_t = 20^o$ and $\theta_e = -35^o$ at the longest periods. Current channelling is important at all periods. According to the errors, which are smaller than in analysis a), a 3D/2D description is more appropriate than a 2D one (Figure 5.14b).

A third analysis c) was performed in this thesis, leaving all parameters and angles free. The strike direction obtained was between -30° and -45° for the short periods (up to 1 s) and around 45° from 1s onward. Twist and shear angles presented sharp variations, particularly the shear. The phases obtained were $\varphi'_{xy} \approx 50^o$, and φ'_{yx} varied from -135° (0.01 s) to 90° (1000 s), reaching greater values at the longest periods (Figure 5.14c).

The decomposition obtained from the phase tensor (Figure 5.10) has, up to 2 s, the same phase values as the 2D description given in analysis a), although with different strike directions. For T>2 s, the phase tensor exhibits 3D effects (β_P different from zero). However, for T>10 s, the strike angles are similar to those obtained from analysis a) and c), but not the phases, since those obtained from the phase tensor are arranged in the first quadrant and do not show if the phases are greater than 90°. As it is expected, the phase tensor and G&B decomposition only lead to comparable results when the data can be described as 2D or 3D/2D. Even in these cases, the phase tensor analysis arranges all the phases in the first quadrant, and it is not possible to detect phases outside of the quadrant, as with the G&B analysis.

(a)
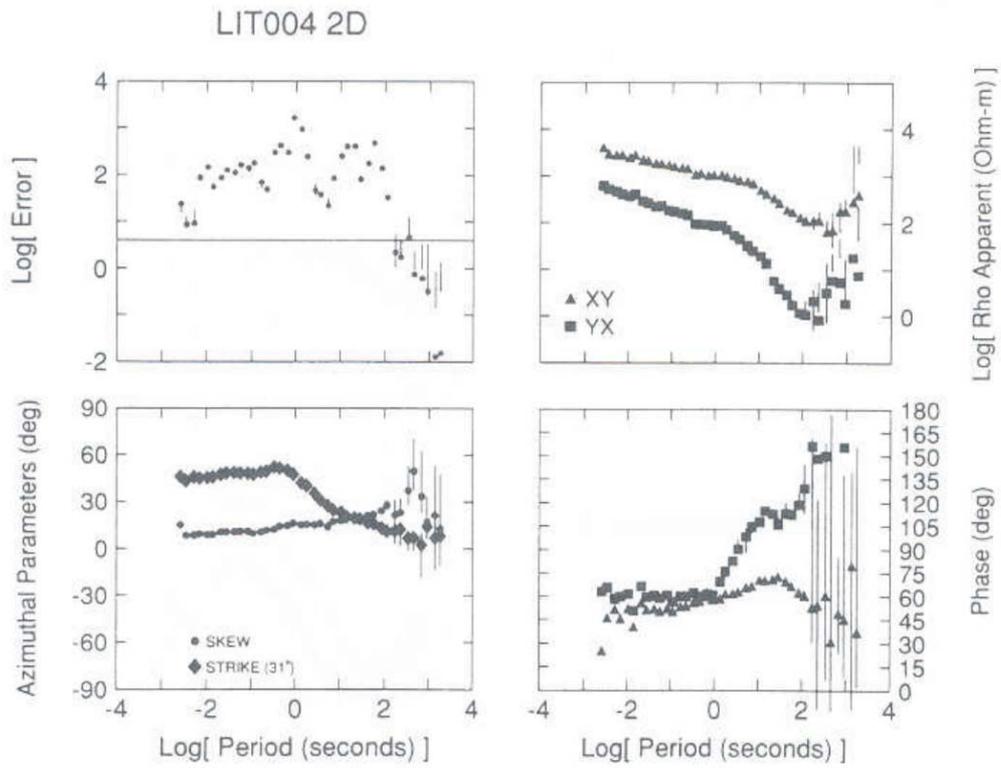
(b)
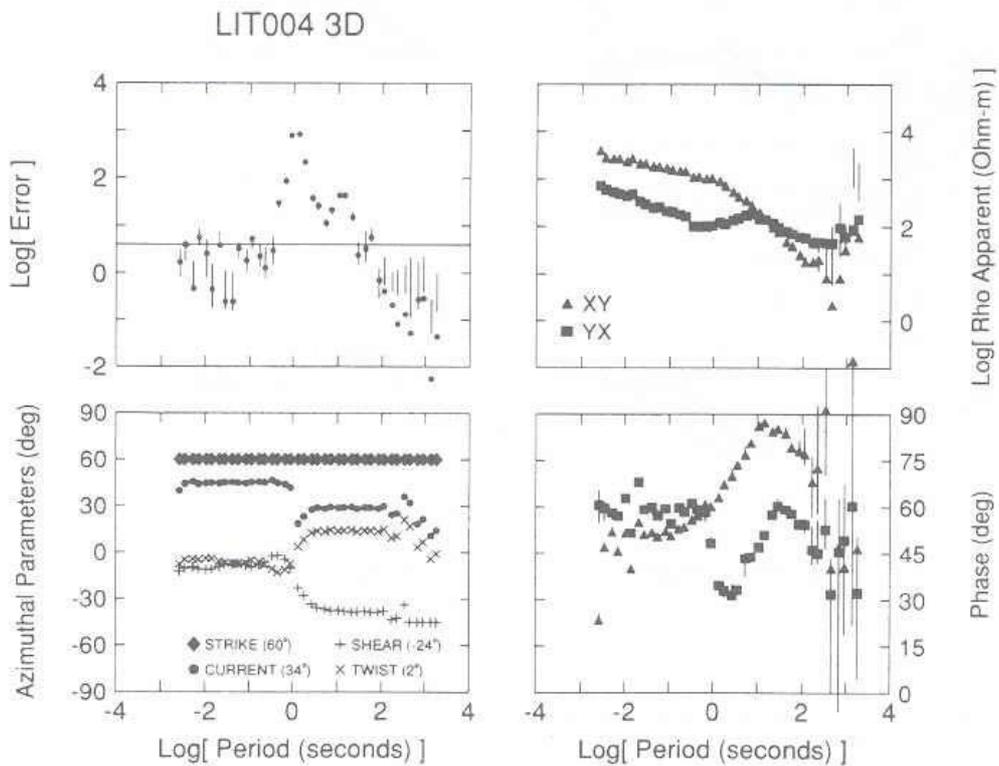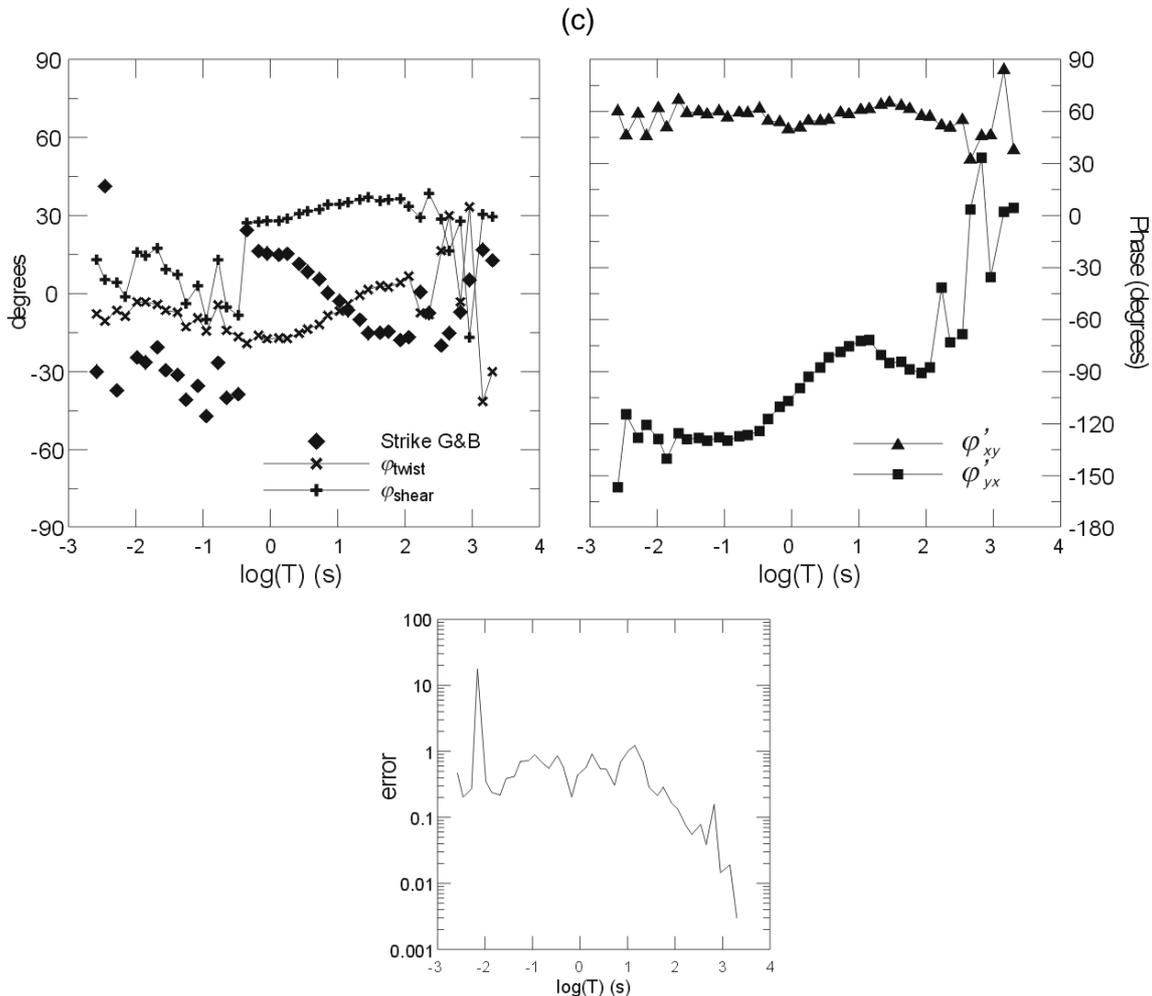
Figure 5.14: Groom and Bailey decomposition parameters using different analysis, corresponding to site 4 from BC87 dataset.   (a): Non-distortion: data rotated to the direction that minimises diagonal components of the MT tensor (Jones *et al.*, 1993). (b): G&B decomposition, allowing galvanic distortion, rotating the data 60° (Jones *et al.*, 1993). (c): G&B decomposition, allowing all parameters free.

5.5 Fitting the Phase Tensor Data to a 2D Model

This section presents an approach to obtain a 2D description of the data, by fitting the measured phase tensor to that corresponding to a 2D model. Contrary to the Multisite Multifrequency Strike code (Strike, McNeice and Jones, 2001), which uses the MT tensor, the proposed approach does not provide distortion decomposition, since the phase tensor does not include this information.

The process of fitting the phase tensor observed data to a 2D phase tensor model was done by minimizing a misfit function between the observed and model parameters of the phase tensor:

$$\gamma^2 = \sum_{k=1}^{N \cdot S} \left(\sum_{i=1}^{4} \left(\frac{\Phi_{iobs} - \Phi_{i\,\mathrm{mod}}}{\sigma_i} \right)^2 \right), \qquad (5.1)$$

where N·S is the number of phase tensors included in the process: N, the number of frequencies and S, the number of sites.

$\Phi_{i\,obs}$ and σ_i (i=1 to 4) are the 4 observed components of the phase tensor and their standard deviations.

$\Phi_{i\,mod}$ (i=1 to 4) are the components that one should obtain for a 2D description of the data, which make a total of 2·N·S+M parameters: N·S·φ_{Max}, N·S·φ_{min} and M·α (different strike directions considered):

$$\Phi_{1\,\mathrm{mod}} = \Phi_{Max} \cdot \cos^2 \alpha + \Phi_{min} \cdot \sin^2 \alpha, \qquad (5.2)$$

$$\Phi_{2\,\mathrm{mod}} = \Phi_{3\,\mathrm{mod}} = \left(-\Phi_{Max} + \Phi_{min} \right) \cdot \cos \alpha \cdot \sin \alpha, \qquad (5.3)$$

$$\Phi_{4\,\mathrm{mod}} = \Phi_{Max} \cdot \sin^2 \alpha + \Phi_{min} \cdot \cos^2 \alpha. \qquad (5.4)$$

Thus, in the case that each site and each frequency could have a different strike direction, the fitting process would imply N·S different functions, each with 4 data and 3 parameters to fit. If all frequencies in every site had a common strike, there would be S different functions to fit, each one with 4·N data and 2·N+1 parameters to fit. Finally, if a common strike was considered for all data, the fitting would use 4·N·S data and there would be 2·N·S+1 parameters to fit.

The minimisation method employed is the same used in the Strike code (McNeice and Jones, 2001), using a Taylor expansion of the γ^2 function with the Jacobian and Hessian matrixes for the 1st and 2nd derivatives.

Furthermore, the fitting process also needs a set of initial parameters, as well as a convergence parameter or, alternatively, a maximum number of iterations to stop the minimisation.

The code created was tested using a subset of the BC87 dataset, referred to as line BC87c. It includes 10 of the western sites, 902, 0, 1, 3, 4, 5, 006, 11, 15 and 17. This subset was chosen because the information from G&B decomposition, using the Strike code with a strike direction of 18° for the period range from 0.01 to 10s, was available for comparison (Jones, pers. comm.). The hypothetical strike values, α_P-β_P, obtained from the phase tensor, compared to the 18° of the G&B decomposition, are displayed in (Figure 5.15), showing broad variations.
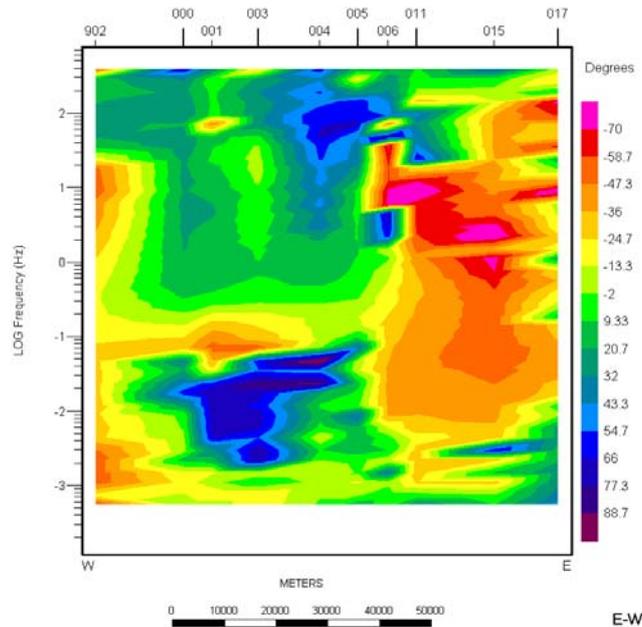
Figure 5.15: α_P-β_P directions obtained along line BC87c from the phase tensor data. ■: 18° (Strike used in the Groom and Bailey decomposition).

A first test was done by trying to adjust all frequencies in each site to the same strike direction. Hence, one function was fitted for each site, with 2·N +1 parameters each. The maximum number of iterations allowed was 100, otherwise, the process stopped when the change in the misfit function between two iterations was less than 10^{-10}. A set of initial parameters was also defined. In general, the final result was independent of the initial parameters. However, in some cases, the final parameters were the same as the initial ones, and any parameter led to the same value of the misfit function, which was notably high. This means that these data cannot, in fact, be adjusted to a 2D model with a unique strike direction. Combining the frequencies into smaller groups made the misfit function smaller, although there were some groups for which the misfit function did not decrease and the iterative process stopped after the 100 iterations.

Three different fitting processes were performed for the BC87c line. The results from site 001, compared to G&B decomposition with an 18° strike, show the following:

1) All parameters free: different strike for every frequency. The aim of this inversion was to check if, even without constraints, there is a predominant strike direction in the final parameters obtained. The strike directions obtained present smooth variations around 0°, but clearly different to the 18° obtained from G&B decomposition. The corresponding phases are similar to those of G&B for the shorter periods. The misfit function is large and converges after 10 iterations.

2) A common strike for every 4 frequencies. The strike directions and phases obtained are similar to those of misfit 1), where these changed smoothly with the period.

3) A common strike for all frequencies. The strike direction obtained was approximately -13°, and the phases similar to 1) and 2).

Although the three inversions led to similar results, these differ significantly from the results of the Groom and Bailey decomposition. For the rest of the sites in this dataset, it was observed that, even if the strike pattern is very different in 1), 2) and 3), the phases are similar, and different to those obtained with Groom and Bailey decomposition.

This small variation of the phases in the strike direction would indicate one-dimensionality, which is not the case. Working only with the phases, and not with the resistivities, is an important limitation. Further tests and research on this minimisation process are necessary to lead to satisfactory results.

5.6 Conclusions

A summary of the phase tensor obtained from synthetic MT tensors has been presented to illustrate its features and graphical representation for different types of dimensionality.

The phase tensors computed from the responses of a synthetic model showed the effects of noise and distortion: the errors lead to wrong interpretations of the dimensionality, and, under extreme distortion conditions, the phase tensor is not distortion-invariant.

The comparison between the phase tensor parameters and WAL invariants allowed the identification of a relationship between φ_{Max} - φ_{min} (difference between the maximum and minimum phases obtained from the phase tensor) and the WAL invariant Q. These magnitudes have the same values when data are not affected by distortion, which introduces a new way to identify distortion even if data are 3D. With regard to the strike angles computed from the phase tensor and from the WAL method using the same data, the errors of the latter are greater.

A comparison of the phases and strike directions obtained from the phase tensor and from the Groom and Bailey decomposition showed significant differences due to the hypothesis of 2D dimensionality that applies in the second, and the use of only the phases information in the phase tensor.

Finally, a code was developed to fit the phase tensor data to a 2D model. The results were quite inconsistent, however, because of few data to work with: 4N data to fit 2N+1 parameters, compared to Groom and Bailey, 8N data and 4N+3 parameters.