To the problem of long-period statics

Georgy Mitrofanov (IPGG&NSU&NSTU, Russia), Viatcheslav Priimenko (LENEP/UENF&INCT-GP/CNPq/MEC, Brazil) and Nikita Goreyavchev (IPGG&NSU, Russia)

Abstract

We consider factor models used in the problems of residual statics and waveform correction. Such models lead to a system of linear equations, which allows us to perform a detailed analysis of the solutions obtained. In particular, it is possible to understand the problem of long-period statics and its connection with acquisition configuration. At the same time, it is shown how this problem can be solved at the algorithmic level with optimization of the used *a priori* information. The latter is especially relevant for the analysis of dynamic problems. In the study of linear systems two approaches are implemented: (i) an iterative process of consistently refining the factor estimates, and (ii) introducing heuristic conditions that ensure uniqueness and stability of the solution.

Introduction

Linear factor models are widely used in statistical analysis, for instance, they are typical in analyzing regression and variance of experimental and field data. One class of these models appeared in seismic exploration. It was originally associated with the tasks of separating waves and temporal statics correction (Holtzman, Troyan 1967; Hileman et al., 1968; Taner et al., 1974). Then these models were extended and generalized to the tasks of analysis and correction of the seismic signal shape (Gurvich, 1970; Goldin, Mitrofanov, 1975; Mitrofanov, 1980; Taner, Koehler, 1981). The essential point for the factor models is their linear structure with respect to the unknown parameters (factors), that allows us to apply the methods of linear algebra and to conduct in-depth studies of the properties of the obtained solutions, see Mitrofanov (1975) and Wiggins et al. (1976).

As a result, non-uniqueness of the solutions was analyzed, as well as the influence on the non-uniqueness of both the factor model and the acquisition configuration. This made it possible to separate the effects affecting the non-uniqueness of determining the structure of the factor model and the observation system. At the first stages, rather short observation profiles were considered, therefore the focus was on rather local features of observations: multiplicity, angular observations. Also, already at that time was a comprehension of the problem of low-frequency statics (Kirkham, Poggnagliolmy, 1976). However, for a long time, this problem, although associated with the structure of observations, was not explained from the point of view of non-uniquely determined components of the model. In our report we will try to give such an explanation. Moreover, it will be shown that this problem is a consequence of some intrinsic links between the factor model and features of the observation system.

These studies were initiated by offshore engineering seismic survey, where have been used the observation systems with significant ratios between profile length and source-receiver spread. The results obtained open up completely new aspects of factor models and may be of interest to researchers in other areas.

Method

Factor models used in seismic surveys to represent plane waves or the residual arrival times of seismic signals, as well as the variations in their shapes, can be given in the following general form (Mitrofanov, 1988)

$$(1) \quad z_{ij} = \sum_{p=1}^{P} \gamma_p \left(c_1^p i + c_2^p j \right) \cdot \varphi_p \left(c_3^p i + c_4^p j \right) + \varepsilon_{ij}.$$

Here functions $\gamma_p \left(c_1^p i + c_2^p j \right)$ and $\varphi_p \left(c_3^p i + c_4^p j \right)$ do not change along lines $c_k^p i + c_{k+1}^p j = \mathrm{const}$, $k = 1,3$. It is this feature that determines the order of parameterization (factorization) of such models. We assume that coefficients c_k^p, $k = 1,2,3,4$, which determine directions or lines for the functions γ_p, φ_p, are known or specified when we define the model; function φ_p is supposed to be known and γ_p - unknown. The task is to determine the values of γ_p or their factors by observed values of z_{ij}, which can be seismograms given in a discrete form or some functional with respect to observed seismic signals, for instance, residual arrival times of signals in the time static correction problem. In the first case, indices i and j determine the trace and time register number, in the second case - the source and receiver numbers in multiple coverage. The additional term ε_{ij} in (1) includes the remaining part of the considered wave field or its functionals, which is not represented by the factor components.

It is obvious that diversity of models (1) is related to P, c_k^p ($k = 1,2,3,4$) and φ_p, and with the structure of the observing system with defined values of z_{ij}. The

obtained estimates are determined by these parameters, and furthermore, their properties of depend on ε_{ij}.

As the most known examples in the seismic exploration, we consider two four-factor models. The first one is a model of residual time correction or static correction:

(2) $z_{ij} = \Delta t_{ij} = \Delta t_i + \Delta t_j + c_{i+j}^0 + c_{i+j}^2 \cdot (i-j)^2 + \xi_{ij}$,

and the second one – a model of surface consistent deconvolution with a fixed frequency ω:

(3) $z_{ij} = \ln\{S(\omega)\} = \alpha_i + \beta_j + \gamma_{i+j} + \lambda_{i-j} + \xi_{ij}$.

These models are based on expression (1) with $c_1^1 = c_2^1 = c_1^2 = 0, c_2^2 = 1$ and $\varphi_p \equiv 1$. Next, for the model (2) we define $c_1^p = c_2^p = 1/2, c_3^p = -c_4^p = 1$ with p=3,4, and select φ_p with even positive powers of the corresponding argument (including zero degree). For model (3) we define $c_1^3 = c_2^3 = c_1^4 = -c_2^4 = 1/2$ and $\varphi_p \equiv 1$. In both models indices i and j correspond to the numbers of sources and receivers, and $i+j$, $i-j$ - to the numbers of the common midpoints and common offsets, respectively. Both models have worked well in seismic data processing. However, the limitations of these models are known. Thus, model (2) requires the symmetry of the residual travel time shifts. This condition is usually provided for monotype waves in the case of profile observing systems. Model (3) does not require symmetry with respect to source-receiver spreading, but only allows to determine general characteristic of the medium at different offsets.

Model (1), as well its particular cases (2)-(3), linearly depends on unknown parameters, which provides a fairly simple study of its properties and can be used in the case of large amounts of input date. But a more significant point is that the obtained estimates of unknown parameters (factors) can be used in solving various problems, for example, accounting waveform variation or determining target parameters of the medium.

The structure of model (1) defines the block type of the corresponding linear system, i.e.,

(4) $z = X\theta + \varepsilon$,

where z, ε consist of values of z_{ij}, ε_{ij}, respectively, ordered in a certain way by observation points, and vector θ consists of certain values of factors γ_p. Matrix $X = \{X_1, X_2, \ldots, X_P\}$, where elements of X_p are defined by the values of φ_p as well as by the representation of a certain value of factor γ_p in fixed z_{ij}. The block structure of X permits to formulate a block linear system, arising in the least square method (LSM), providing a solution to overdetermined system (4)

when the number of observations is greater than the number of unknown factors. It also allows us to formulate an iterative process in the following matrix form (Mitrofanov, 1978):

(5) $\theta^{(l)} = T\theta^{(l-1)} + B\hat{z}$,

where

(6) $\begin{cases} T = -\sum\limits_{j=1}^{P-1} C_j^P E_j R_{j+1\ldots P}^j, \\ B = E_p + \sum\limits_{j=1}^{P-1} C_j^P E_j, \\ C_j^P = \prod\limits_{i=P}^{j+1} (I_{i-1} - E_i R_{j+1\ldots i-1}^j) \end{cases}$.

Here $\theta^{(l)}$ is the vector of factor estimates obtained at the l-iteration, and $\hat{z} = X^T z$. Matrixes $E_j, R_{k\ldots m}^j$ are composed of $X_k^T X_j$ or $(X_j^T X_j)^{-1}$, and I_j are blocks of the identity matrix. In this case, matrixes $X_k^T X_j$ coincide with the blocks of the matrix of the LSM-system.

Iterative process (5) is a particular case of the over-relaxation method (Marchuk and Kuznetcov, 1975), which converges to the solution of system (4) obtained by the least square method. The method does not require the construction of a common matrix and can be simply organized, which is essential in the case of a large amount of the data. It also weakly depends on the possible cut of few observations under their analysis. At the same time, on its basis it is possible to investigate the structure of non-uniquely determined factor components. The latter is important when choosing the required *a priori* information.

When the question of the structure of non-uniquely determined factor components is resolved, then the *a priori* information, providing a unique solution of system (4), can be formed. The type of *a priori* information is determined by field data or some heuristic ideas that may be, for instance, of finite difference or statistical nature. For factor models (1) such ideas are rather simply formulated in a linear form (Scheffe, 1980; Mitrofanov, 1988), which makes it possible to consider a generalized version of the LSM-system (using Lagrange's undefinite multipliers), see (Rao, 1968)

(7) $\begin{cases} X^T X \theta_m = \hat{z} \\ H\theta_m = c \end{cases}$,

where H, c are determined by the type of *a priori* data. In the absence of real *a priori* information, the second matrix equation can be formed on the basis of the indicated heuristic ideas. In this case, we usually have $c = 0$. The obtained solution can be corrected using *a priori* information having an optimum amount of data and form (Mitrofanov et al., 2011).

Examples

We utilized the marine data obtained using submerged receivers in the White See, the north part of Russia. The data had a wide frequency spectrum and required high-quality processing. The profile of observations was made up of 6231 sources configured with 16 receivers. It was necessary to analyze the variability of the form of well-allocated direct wave, taking into account the deeper receiver's system. For such an analysis, the following two-factor model was used:

$$(8) \quad z_{ij} = \gamma_1(i) + \gamma_2(j) + \varepsilon_{ij} \, ,$$

which was obtained from (1) with $P = 2$, $c_1^1 = c_2^2 = 1$,

$c_2^1 = c_1^2 = 0$, and $\varphi_p(c_3^p i + c_4^p j) \equiv 0$. Obviously, it will be equivalent to the truncated versions of models (2)-(3), when they contain only factors: $\Delta t_i, \Delta t_j$ or α_i, β_j. It is also equivalent to the classical two-factor model used in the dispersion analysis (Scheffe, 1980). The model is well studied, and its non-uniqueness is connected only with the constant component. But it turned out that this completely describes the problem of non-uniqueness for the cases of surface observation systems used in three-dimensional seismic exploration. For the observation systems, used in seismic prospecting and significantly different from the classical models of dispersion analysis, the problem becomes more complex.

In order to analyze the influence of the observation system on the obtained solution, we conducted several experiments. Below are results related to the experiments in which variations of factors for sources and receivers were simulated by sinusoidal functions with different frequencies. The experiments are related to the well-known problem of low-frequency or long-period statics. The meaning of the problem lies in the fact that when we estimate the factor parameters related to sources and receivers in a real seismic experiment, only high-frequency variations (with a period shorter than the length of arrangement) are well estimated.

Fig.1 shows two observation systems containing 160 sources. The left system has 128 receivers in the arrangement, and the right one - 16. As a result, for both observation systems there were 160 values of the source factor $\gamma_1(i)$ and a different number of values for the receiver factor $\gamma_2(j)$. For the left system such values were 287, and for the right one - 175.

For both systems, model data were generated as a sum of factor values for the source and receiver. The values are represented by sinusoid segments. They are shown in Fig.2(a): blue color - the factor values for source, and the red - for receiver. Using model data, an iterative process was performed to determine the values of the factors. Fig.2(b) shows the obtained results. At the same time, the top row is related to the observation system containing 128 receivers, and the bottom row - 16 receivers. Fig.2(c) shows the logarithm of the standard deviation calculated after each iteration.

According to the results presented, for the observation system with 128 receivers in the spread, it was possible to determine accuracy the values of parameters with high convergence speed of the iterative process. For the second observation system, where the period of parameter variations is several times the length of the receiver arrangement, the convergence speed of the iterative process was slowed down considerably, and the obtained estimates were low accuracy. It can be noted that in the factor values it is possible to separate the spatial frequencies, but some trend components having a periodic structure begin to appear.

Due to the fact that the real observation system contained the 16-channel arrangement with a significant number of sources, we carried out experiments, which have been focused precisely on this type of observation systems. Figs.3 shows the results for some of these experiments. The top figure shows the factor values, the middle part shows the results of their determination, and the bottom one – the logarithm of the standard deviation by iterations. The results characterize the possibility of separating the factors for different spatial frequencies, presented in the observations and associated with waveform variations in the sources and receivers. The carried-out experiments also made it possible to formulate a heuristic idea about the nature of long-period components, which can be formed under evaluating the factor estimates.

Thus, the factors are stably defined with increasing spatial frequencies for the observing system with 16 receivers in array. This is shown in Fig.3(a) and Fig.3(b). We can see that in the case of variations in the original factors with a period of about the length of the array, we had a good quality estimates of factors, see Fig.3(b). However, with increasing the period up to 1,5 array length, and, respectively, the period of receiver factor variations up to 3 array length, there is deterioration in the quality of the determined parameters. At that a noticeable correlation of the source factors variation with slower receiver factors changes have been observed, see Fig.3(a). A similar dependence also occurs for systems containing a larger number of sources. Moreover, this feature will continue in the case of more complex factor models. As an example, in Fig. 4 presents the results of the evaluation of factors for model (3) with $\lambda_{i-j} \equiv 0$.

This correlation shows that in the process of sequential refinement for systems with a moving arrangement and a large number of sources (many times exceeding the length of seismic array), the ambiguously determined components of the factor model are localized. In particular, for two-factor model (8), where only the constant component is non-uniquely determined, such localization leads to a redistribution of its value along the profile.

As a result, we have the ambiguity of determining low-frequency components of the factors, if they are present simultaneously, both in sources and in receivers. This non-uniqueness is determined by the observation system, which does not provide enough coherence between the various observations and the model parameters.

This coherence can be provided by the conditions introduced into system (7) in the form of a pseudo *a priori* information (Mitrofanov, 1988). For two-factor models (8), this condition may be quite simple, for instance,

$$(9) \quad \sum_{j=1}^{J} \gamma_2(j) = 0$$

which relates the constant component to the source factor. Such conditions provide (for model data) the accuracy separation of factors through a solution of (7), as evidenced by the results in Fig.5.

The figure shows the results of the separation of factor values in the form of two units (Fig.5(a)) and low-frequency sinusoidal waves (Fig.5(b), with a period exceeding the length of the 16-channel array 1.5 times, i.e., this figure corresponded to the experiment shown in Fig.3(a). Small shifts in the values of certain factors are associated with the redistribution of the constant component from the value of receiver factor to the values of the source factor in accordance with (9). But such redistribution is not local, but global, leading to a general shift in all values of the factor. As a result, we are freed from the problem of long-period static, caused by the feature of observation system with moving receivers. The same results we will get for factors models with the complex structure, see Fig.6.

Conclusions

The results obtained allow us to understand how significant observation system and the estimation method affect the quality of factors separation. At the same time, a significant number of experiments performed showed that for observation systems with a large number of sources (many times exceeding the length of seismic array) the non-uniquely determined component of the factor model is localized. Such localization leads to redistribution of its value along of profile. As a result, we have the non-uniquely defined low-frequency components, if they are presented simultaneously in both sources and receivers. Such non-uniqueness is defined by the system of observations, which does not provide sufficient coherence between observations and the model parameters. Understanding this fact allows us to provide the mentioned above coherence with the conditions introduced into the extended linear system (9). In this case, the problem of long-period static, which is determined by the ratio of the length of the profile and the seismic array, is eliminated.

Acknowledgments

References

GOLTCMAN, F.M. and TROYAN, V.N. Optimal algorithm for the separation of interfering seismic waves. Izv. AN SSSR, Physics of Earth, V.8, p.101-112, 1967. (in Russian)

GOLDIN, S.V. and MITROFANOV, G.M. Spectral and statistical method of accounting of surface irregularities in the multiple observation systems of reflected waves. Soviet Geology and Geophysics, V.6, p.102-112, 1975. (in Russian)

GURVICH, I.I. On the theoretical basis of dynamic measurements in the seismic survey. Izv. Vuzov, Geology and Exploration, V.6, p.108-113, 1970. (in Russian)

HILEMAN, J.A., EMBREE, P., PFLUEGER J.C. Automated static corrections. Geophysical Prospecting, V.18(3), 1968.

KIRKHAM, D.S. and POGGNAGLIOLMY, E. Long period statics determination by inverse filtering. Geophys. Prospect., V.24(4), 1976.

MARCHUK, G.I. and KUZNECOV, Yu.A. Iterative methods and quadratic functionals. In: Methods of Computational Mathematics. Novosibirsk, Nauka, 1975. (in Russian).

MITROFANOV, G.M. Joint estimate of linear factors in the CDP observation systems. In: Using method of numerical mathematics and mathematical statistics at digital processing of seismic data. Novosibirsk, IPPG SO RAN, 1975. (in Russian)

MITROFANOV, G.M. Effective representation of the wave field in seismic exploration. Soviet Geology and Geophysics, V.4, p.135-145, 1980.

MITROFANOV, G.M. Pseudo a priori information in the problem of correcting the frequency-dependent static. In: Mathematical Problems of Seismic Data Interpretation. Novosibirsk, Nauka, 149-168, 1988. (in Russian)

MITROFANOV, G., PRIIMENKO, V., MISSÁGIA, R. Utilização efetiva da informação a priori na inversão da forma de sinal sísmico. Revista Brasileira de Geofísica, V.29(2), p.247-263, 2011.

RAO, S.R. Linear Statistical Inference and their Applications, 2ed. New York, John Wiley & Sons, Inc., 1973.

SCHEFFE, H. The Analysis of Variance. New York, John Wiley & Sons, Inc., 1959.

TANER, M.T. and KOEHLER, F. Surface consistent corrections. Geophysics, V.46(1), p.17–22, 1981.

TANER, M.T., KOEHLER, F., ALHILALI, K.A. Estimation and correction of near-surface time anomalies. Geophysics, V.39(4), 1974.

WIGGINS, R.A., LARNER, K.L., WISECUP, R.D. Residual analysis as a general linear inverse problem. Geophysics, V.41(5), 1976.
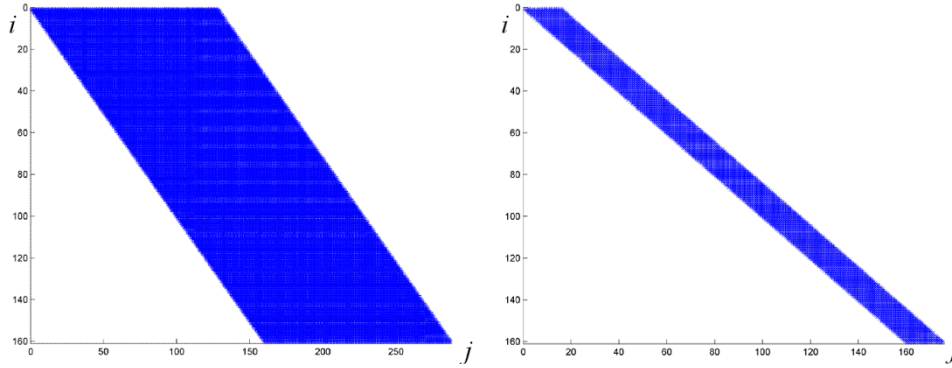
Figure 1. Observation systems with different number of receivers in array.
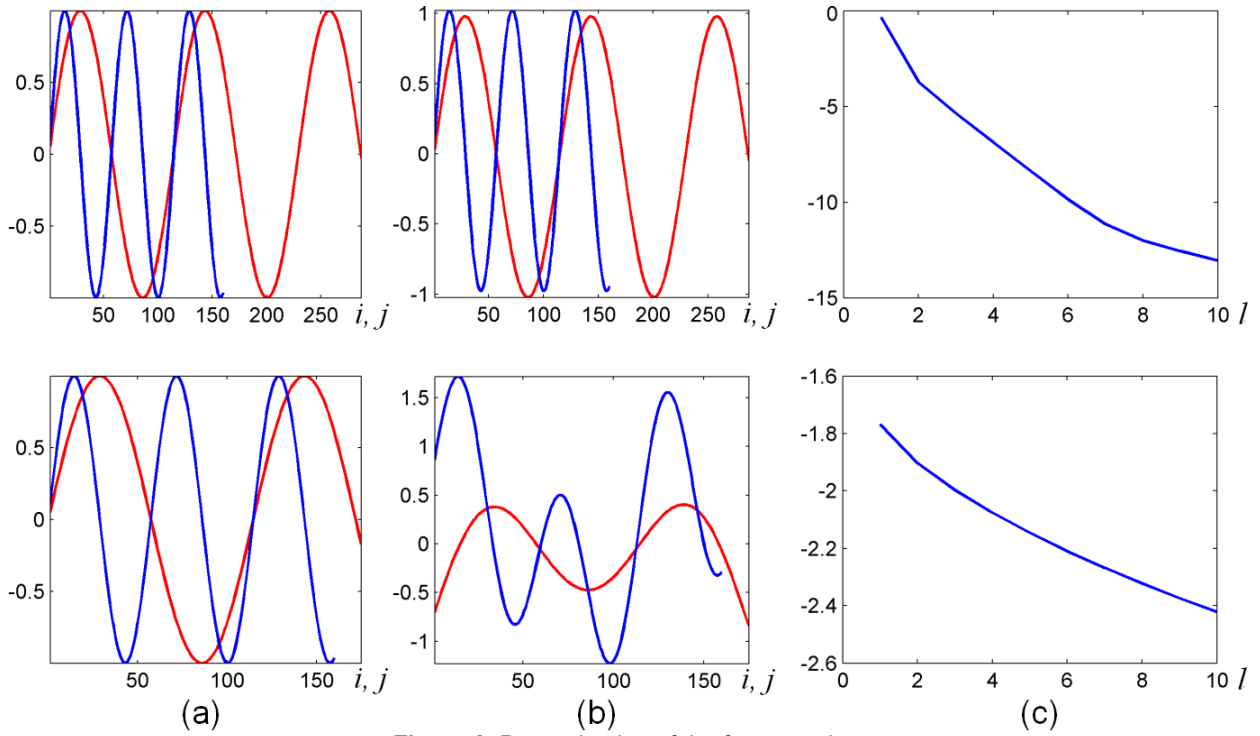

(a)   (b)   (c)

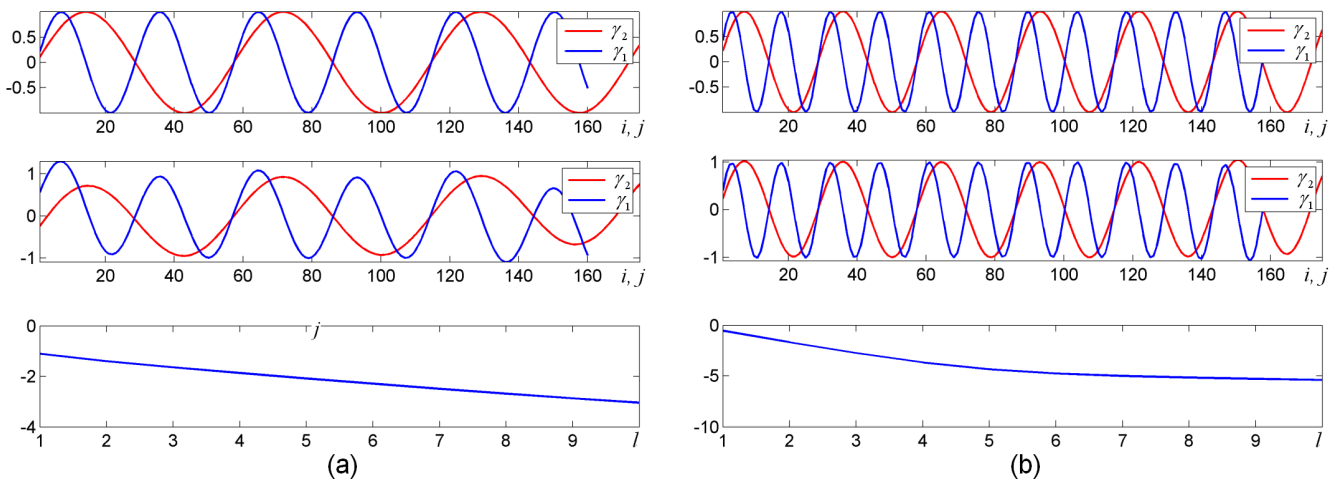**Figure 2**. Determination of the factor vaules.

(a)   (b)

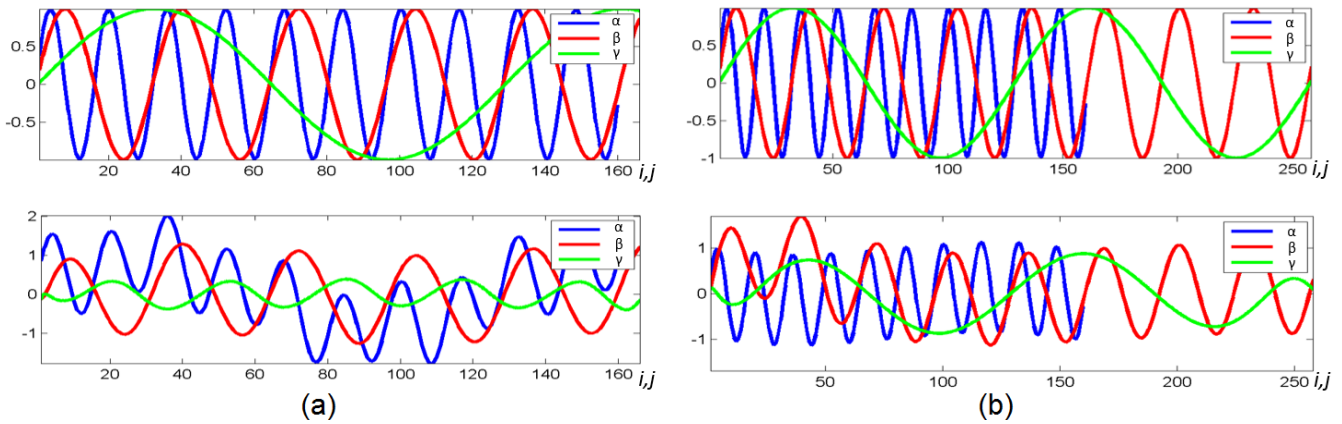**Figure 3.** Determination of the factor values with different spatial frequency for observations with 16 receivers in array.

Figure 4. Determination of the factor values in more complex model using observations with different number of receivers in array: (a) 16 and (b) 128.
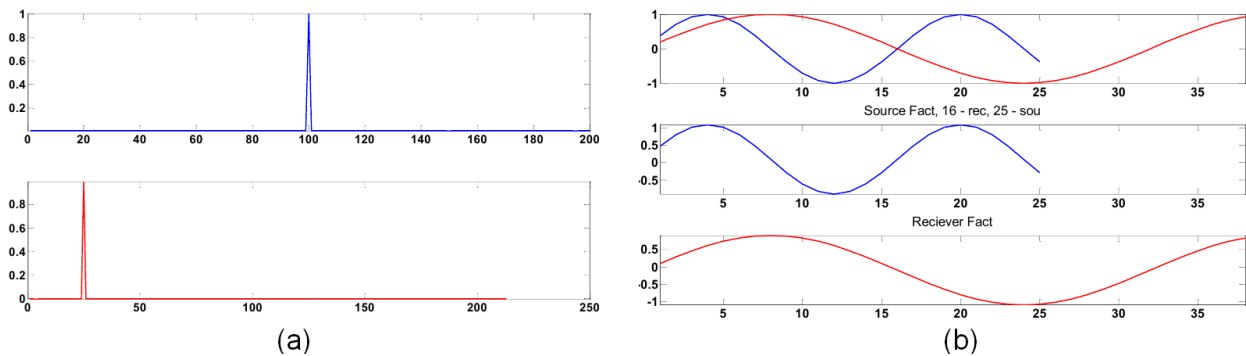

Figure 5. Determination of the factors using the solution of linear system (7).
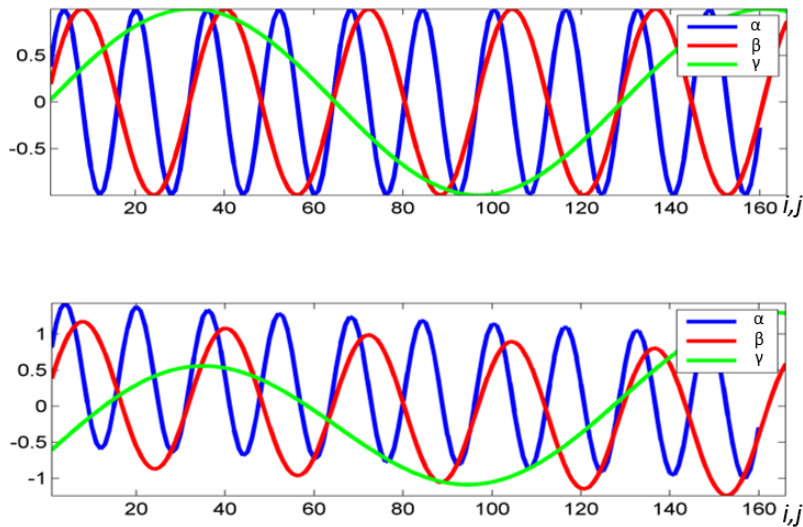

Figure 6. Determination of the factors values in more complex model using the solution to extended linear system (7).