The preconditioned biconjugate gradient algorithm applied to geophysical electromagnetic modeling

Victor Cezar T. de Souza, Luiz Rijo and Marcos Welby. C. Silva, UFPA, Brazil

Abstract

Solution of the large system of equations that results from a finite-element approximation in electromagnetic modelling over a mesh strongly refined is most commonly done iteratively, because direct solution requires excessive amounts of memory and computation. The conjugate gradient – type methods used in combination with preconditioning – are among the most effective iterative procedures for solving large sparse nonsingular systems of linear equations $\underset{\sim}{\mathbf{A}}\mathbf{x} = \mathbf{b}$. Three test models are performed by the proposed algorithm. The first is the modeling of the geoelectric field under the Equatorial Electrojet employing the interpretative model of Parnaíba Basin Conductivity Anomaly (Silva and Rijo, 2003), the second the electromagnetic tomography, applicable to cylindrical with azimuthally symmetric geometry about vertical magnetic dipole sources (Souza and Rijo, 2003). The third test model is performed using the MCSEM – marine controlled-source electromagnetic (Ellingsrud, 2002) using a 2-D approximation to the sub-seafloor structure. The performance of the algorithm is showed in contrast with the Gaussian elimination.

Introduction

Solution of the large system of equations that results from a finite-element approximation (FE) over a mesh strongly refined is most commonly done iteratively, because direct solution requires excessive amounts of memory and computation. Although some progress has been made in applying algorithms for the direct solution in banded systems to FE equations, these still require hours of computational time.

In this paper, the biconjugate gradient method is applied with preconditioning to the solution of systems of equations that result from finite element approximations in electromagnetic modeling (se e.g. Coggon, 1971, Rijo, 1977 and Pridmore et al. 1981). The method will be described, with a brief discussion of techniques of preconditioning; it will be performed by three test problems, namely, the MT modeling to Equatorial Electrojet (Silva and Rijo, 2003), electromagnetic tomography (Souza and Rijo, 2003) and the marine controlled-source electromagnetic (MCSEM). Computational time statistics will be given for the test problems analyzing the performance of the method proposed.

The biconjugate algorithm

Conjugate gradient – type methods used in combination with preconditioning – are among the most effective iterative procedures for solving large sparse nonsingular systems of linear equations

$$\underset{\sim}{\mathbf{A}}\mathbf{x} = \mathbf{b} \tag{1}$$

The archetype of these schemes is the classical conjugate gradient algorithm (CG hereafter) of Hestenes and Stiefel, 1952, used for Hermitian positive definite matrices $\underset{\sim}{\mathbf{A}}$. The method can be seen as an iterative process to solve linear equation by minimizing quadratic functional over certain spaces called Krylov spaces (Axelsson, 2000). Their application in electromagnetic methods range to solve linear systems in forward modeling until parameter technique inferring in inverse problem (see, e.g., Zhang et al., 1995 and Wu et al., 2003).

While most linear systems that arise in practice have real coefficient matrices $\underset{\sim}{\mathbf{A}}$ and real right-hand sides $\mathbf{b}$. In numerical modeling of electromagnetic problems are involved complex coefficient functions and/or complex boundary conditions (see, e.g., Coggon, 1971). Hence, the coefficient matrices $\underset{\sim}{\mathbf{A}}$ that arise from theses problems are complex symmetric and non-Hermitian (Freund, 1992). Since the matrix $\underset{\sim}{\mathbf{A}}$ is not Hermitian symmetric, it become one matrix $\underset{\sim}{\mathbf{A}}$ ill-conditioned, and the standard conjugate gradient method cannot be directly applied to solve the system $\underset{\sim}{\mathbf{A}}\mathbf{x} = \mathbf{b}$ (Freund, 1992). On the other hand, the system could be transformed to a Hermitian system by forming the normal equations $\underset{\sim}{\mathbf{A}}^{\mathbf{h}}\underset{\sim}{\mathbf{A}}\mathbf{x} = \underset{\sim}{\mathbf{A}}^{\mathbf{h}}\mathbf{b}$, where $\underset{\sim}{\mathbf{A}}^{\mathbf{h}}$ denotes the conjugate transpose of $\underset{\sim}{\mathbf{A}}$. However, if the matrix $\underset{\sim}{\mathbf{A}}$ is ill-conditioned, the matrix $\underset{\sim}{\mathbf{A}}^{\mathbf{h}}\underset{\sim}{\mathbf{A}}$ is much more ill-conditioned, leading to slow convergence rates for conjugate gradient technique.

The biconjugate gradient method is a generalization of conjugate gradient method that arises from Lanczos's extension of $\underset{\sim}{\mathbf{A}}\mathbf{x} = \mathbf{b}$ to the Hermitian symmetric system.

$$\begin{bmatrix} \mathbf{0} & \underset{\sim}{\mathbf{A}} \\ \underset{\sim}{\mathbf{A}}^{\mathbf{h}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{x}} \\ \mathbf{x} \end{bmatrix} = \begin{bmatrix} \mathbf{b} \\ \tilde{\mathbf{b}} \end{bmatrix} \tag{2}$$

The vector $\tilde{\mathbf{b}}$ which is used to extend $\mathbf{b}$ is chosen as a matter of convenience and determines the value of the extension to the unknown vector $\mathbf{x}$. This avoids forming the poorly conditioned normal equations (Smith, 1996). For symmetric complex matrices $\underset{\sim}{\mathbf{A}}$ the choice of $\tilde{\mathbf{b}}$ made in Jacobs (1981) results in symmetries that reduce the computation needed for one iteration. The resulting method is the same as the conjugate gradient algorithm for complex matrices, with all conjugate transposes replaced with simple transposes. This technique has been discussed in works of Jacobs (1981) and van der Vorst (2000) and in the context of electromagnetic modeling by Sarkar, (1987) and Smith et al. (1990). An iteration of the algorithm is given by

$$\begin{aligned}
\alpha_i &= \frac{\mathbf{r}_i^t \mathbf{w}_i}{\mathbf{p}_i^t \underset{\sim}{\mathbf{A}} \mathbf{p}_i}, \\
\mathbf{x}_{i+1} &= \mathbf{x}_i + \alpha_i \mathbf{p}_i, \\
\mathbf{r}_{i+1} &= \mathbf{b} + \underset{\sim}{\mathbf{A}} \mathbf{x}_{i+1} = \mathbf{r}_i + \alpha_i \underset{\sim}{\mathbf{A}} \mathbf{p}_i, \\
\mathbf{w}_{i+1} &= \mathbf{r}_{i+1}, \\
\beta_{i+1} &= \frac{\mathbf{r}_{i+1}^t \mathbf{w}_{i+1}}{\mathbf{r}_i^t \mathbf{w}_i}, \\
\mathbf{p}_{i+1} &= \mathbf{w}_{i+1} + \beta_{i+1} \mathbf{p}_i,
\end{aligned} \tag{3}$$

where $\mathbf{x}_i$ is the current approximate solution, $\mathbf{r}_i$ is the current residual, $\mathbf{p}_i$ is the search direction in which the approximate solution is altered and $\mathbf{w}_{i+1}$ will be redefined below. For an initial guess of $\mathbf{x}, \mathbf{x}_g$ the standard initializations are

$$\mathbf{x}_0 = \mathbf{x}_g, \ \ \mathbf{r}_0 = \beta_0 - \underset{\sim}{\mathbf{A}} \mathbf{x}_0, \quad \mathbf{w}_0 = \mathbf{r}_0, \ \mathbf{p}_0 = \mathbf{r}_0, \tag{4}$$

The preconditioning

Meijerink and van der Vorst (1977) showed that the conjugate gradient method can be accelerated greatly by using an approximate factorization of $\mathbf{A}$

$$\underset{\sim}{\tilde{\mathbf{C}}}^t \underset{\sim}{\tilde{\mathbf{C}}} \approx \underset{\sim}{\mathbf{A}} \tag{5}$$

to precondition the system $\underset{\sim}{\mathbf{A}}\mathbf{x} = \mathbf{b}$. Preconditioning effectively changes the problem being solved to a better conditioned system

$$\underset{\sim}{\tilde{\mathbf{C}}}^{-t} \underset{\sim}{\tilde{\mathbf{A}}} \underset{\sim}{\tilde{\mathbf{C}}}^{-1} \mathbf{x}' = \underset{\sim}{\tilde{\mathbf{C}}}^{-t} \mathbf{b}_1 \tag{6}$$

where $\mathbf{x}' \equiv \underset{\sim}{\tilde{\mathbf{C}}}\mathbf{x}$, and $\underset{\sim}{\tilde{\mathbf{C}}}^t \equiv \left(\underset{\sim}{\tilde{\mathbf{C}}}^t\right)^{-1}$. After eliminating $\mathbf{w}$ from equations the biconjugate gradient method can be applied to the preconditioned system by making the substitutions:

$$\underset{\sim}{\mathbf{A}} \to \underset{\sim}{\tilde{\mathbf{C}}}^{-t} \underset{\sim}{\mathbf{A}} \underset{\sim}{\tilde{\mathbf{C}}}^{-1}, \ \mathbf{x} \to \underset{\sim}{\tilde{\mathbf{C}}}\mathbf{x}, \ \mathbf{b} \to \underset{\sim}{\tilde{\mathbf{C}}}^{-t}\mathbf{b}, \ \mathbf{r} \to \underset{\sim}{\tilde{\mathbf{C}}}^{-t}\mathbf{r}, \ \mathbf{p} \to \underset{\sim}{\tilde{\mathbf{C}}}\mathbf{p},$$

With a little algebra this method reduces to equations (3) but with $\mathbf{w}_{i+1}$ redefined as

$$\mathbf{w}_{i+1} = \underset{\sim}{\tilde{\mathbf{A}}}^{-1} \mathbf{r}_{i+1} \tag{7}$$

where $\underset{\sim}{\tilde{\mathbf{A}}}^{-1} \equiv \underset{\sim}{\tilde{\mathbf{C}}}^{-1} \underset{\sim}{\tilde{\mathbf{C}}}^t$.

This preconditioning is necessary, because for complex systems the convergence is not guaranteed. Hence all the algorithms in practice use some kind of preconditioning. The preconditioner most used in linear systems derived of electromagnetic problem are: the (symmetric successive overrelaxation (SSOR) and the based in incomplete ILU factorizations, like, the preconditioner incomplete Cholesky with its variants: the modified incomplete Cholesky (MIC) and the shifted incomplete Cholesky (SIC) (see e.g. Wu, 2003). We used as preconditioning the ILU0 factorization (Benzi, 2002). In this technique, the Gaussian elimination is performed without the fill-in, namely, no fill-in is permitted, elsewhere in positions in that the elements are non-zeros. Hence the computational memory spent to storage the preconditioning matrix is the same allocated for the coefficient matrix. The version preconditioned for the biconjugate gradient algorithm (Axelsson and Barker, 1984) is given by

$$\begin{aligned}
\alpha_i &= \frac{\mathbf{r}_i^t \mathbf{w}_i}{\mathbf{p}_i^t \underset{\sim}{\mathbf{A}} \mathbf{p}_i}, \\
\mathbf{x}_{i+1} &= \mathbf{x}_i + \alpha_i \mathbf{p}_i, \\
\mathbf{r}_{i+1} &= \mathbf{b} + \underset{\sim}{\mathbf{A}} \mathbf{x}_{i+1} = \mathbf{r}_i + \alpha_i \underset{\sim}{\mathbf{A}} \mathbf{p}_i, \\
\mathbf{w}_{i+1} &= \underset{\sim}{\tilde{\mathbf{A}}}^{-1} \mathbf{r}_{i+1}, \\
\beta_{i+1} &= \frac{\mathbf{r}_{i+1}^t \mathbf{w}_{i+1}}{\mathbf{r}_i^t \mathbf{w}_i}, \\
\mathbf{p}_{i+1} &= \mathbf{w}_{i+1} + \beta_{i+1} \mathbf{p}_i,
\end{aligned} \tag{8}$$

The storage of matrix A

The application of the finite element method in electromagnetic modeling lead up arisen of sparse matrices. In order to take advantage of the large number of zero elements of such matrices, special schemes are required to represent so far as possible only the nonzero elements, and to be able to perform the common matrix operations. There are many methods for the storing the data that represent such matrices, like the compressed row storage (CRS), compressed and its column version, compressed column storage (CCS), the storing by diagonal called compressed diagonal storage (CDS) etc. (see for instance Saad, 2000 and Barret, M. et al. 1993). The scheme which is used in this work is the so-called Ellpack-Itpack format (Saad, 1996) wich is popular on vector machines. The assumption in this scheme is that there are at most Nd nonzero elements per row, where Nd is small. Then two rectangular arrays of dimension $(Nd \times n)$ each are required (one complex and one integer). The first array, COEF contains the nonzero elements of matrix. The nonzero elements of each row of the matrix can be stored in a column of the array COEF(1:Nd,1:n), completing the column by zeros as necessary. Together with COEF, an integer array JCOEF(1:Nd,1:n) must be stored which contains the row positions of each entry in COEF. One example, this scheme is represented follow-up in Figure 1, by one typical symmetric matrix that emerges of the finite element modeling.

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & 0 & a_{14} & 0 & 0 \\ a_{12} & a_{22} & a_{23} & a_{24} & a_{25} & 0 \\ 0 & a_{23} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{14} & a_{24} & a_{34} & a_{44} & a_{45} & 0 \\ 0 & a_{25} & a_{35} & a_{45} & a_{55} & a_{56} \\ 0 & 0 & a_{36} & 0 & a_{56} & a_{66} \end{bmatrix}$$

$$\text{COEF} = \begin{bmatrix} a_{11} & a_{22} & a_{33} & a_{44} & a_{55} & a_{66} \\ a_{12} & a_{23} & a_{34} & a_{45} & a_{56} & 0 \\ a_{14} & a_{24} & a_{35} & 0 & 0 & 0 \\ 0 & a_{25} & a_{36} & 0 & 0 & 0 \end{bmatrix}$$

$$\text{JCOEF} = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 4 & 5 & 6 & 0 \\ 4 & 4 & 5 & 0 & 0 & 0 \\ 0 & 5 & 6 & 0 & 0 & 0 \end{bmatrix}$$

Figure 1 – The Ellpack-Itpack scheme for storage sparse matrices that emerge of numerical modeling by finite element method.

Test models performed by BCG algorithm

Three test models are performed by the proposed algorithm. The first is the modeling of the geoelectric field under the Equatorial Electrojet employing the interpretative model of Parnaíba Basin Conductivity Anomaly (Silva and Rijo, 2003), as is shown in Figure 2. The field electric obtained are introduced in Figure 6.

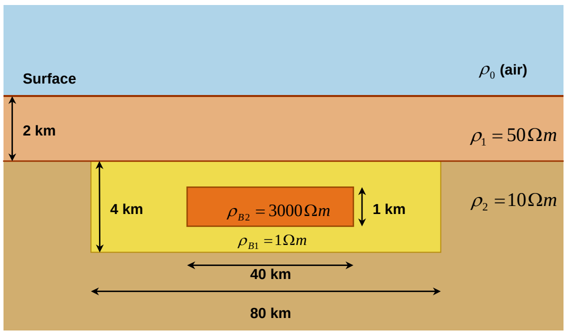

Figure 2 – Model of Parnaíba Basin Conductivity Anomaly.

The second test problem is the electromagnetic tomography, applicable to cylindrical with azimuthally symmetric geometry about vertical magnetic dipole sources (Souza and Rijo, 2003). The model consists of bodies anomalous in an otherwise homogeneous background of electrical conductivity σ^p. Vertical magnetic dipole sources are laid upon the symmetric axis that represents a well. One example of this geometry is the classical model introduced by Alumbaugh, (Alumbaugh and Morrison, 1995). The Figure 3 illustrates four targets taking shape of the acronym **SbgF**. The Figure 6 illustrates the results obtained by tomography EM.

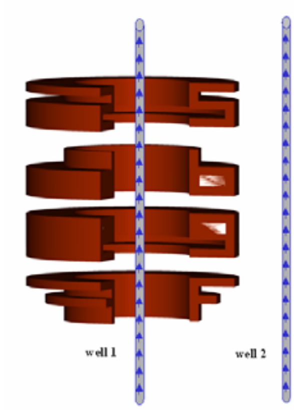

Figure 3 – Three-dimensional representation of complex geometry. A section of anomalous bodies between the wells show us the acronym – **SbgF**

The last test model is performed using the MCSEM – marine controlled-source electromagnetic (Ellingsrud, 2002) using a 2-D approximation to the sub-seafloor structure, as introduces in the Figure 4. In this sketch the hydrocarbon bearing has thickness h_3 and the reservoir is located in a depth h_2 embedded at halfspace beneath the seafloor. The model is energized by a mobile horizontal electric dipole (HED) source and an array of seafloor electric field receivers.

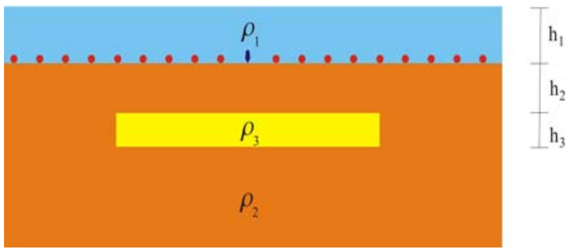

Figure 4 – Sketch of the model. Red marks indicate receivers, blue mark the transmitter.

The Figure 7 illustrates the electrical fields for a model, in which the resistivity of the sea is equal to 0.3 ohm-m, and a level of 800 m for the seawater is considered. The resistivity of the host is of 1 ohm-m, and the reservoir possesses resistivity of 100 ohm-m and strikes at depth of 1Km with a thickness of 100m (Eidesmo, et..al., 2000). The normalized electric field strength is showed in Figure 8 as function of range for the in-line geometry. The other model is introduced in Figure 9; in this case the electric fields are very near each other, because the model is more demanding, which the follows values: ρ_1 = 0.3 Ωm, ρ_2=1 Ωm and ρ_3 = 10 Ωm, the height are: h1 = 1500 m , h2 = 2450 m and h3 = 10 m. The Figure 10 shows the normalized in-line electrical field resulting from this model.

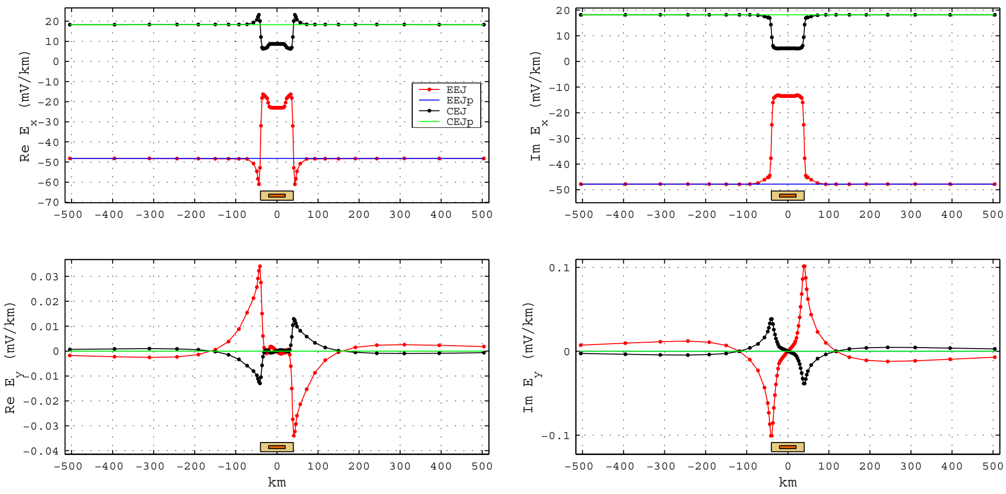

Figure 5 – Geoelectric field components in the presence of 2-D structure under the Equatorial Electrojet.

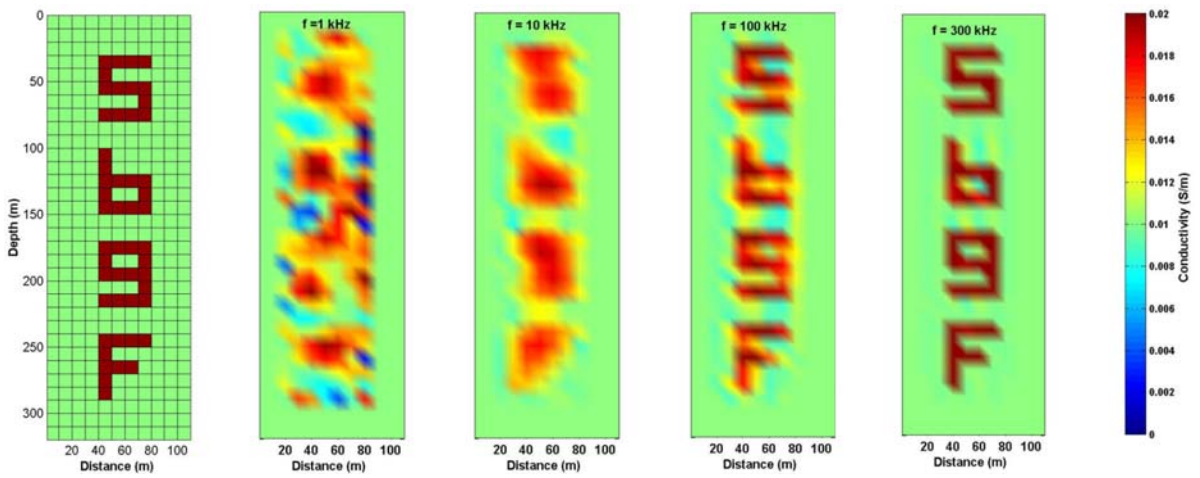

Figure 6 - Results for anomalous bodies with complex geometries four frequencies. (a) True model, - recovered image - at: (b) 1 kHz, (c) 10 kHz, (d) 100 kHz and (f) 300 kHz.

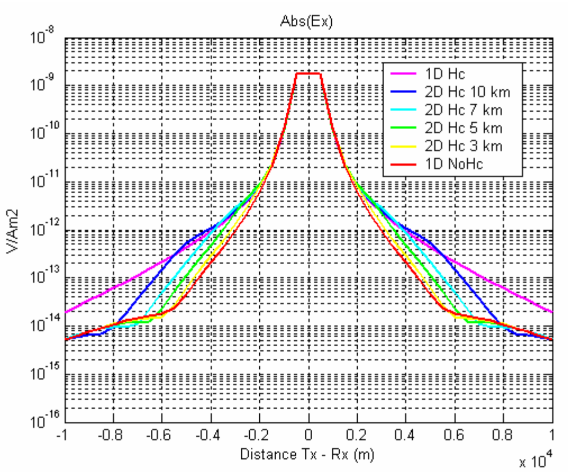

Figure 7 - Amplitude of the in-line electrical field

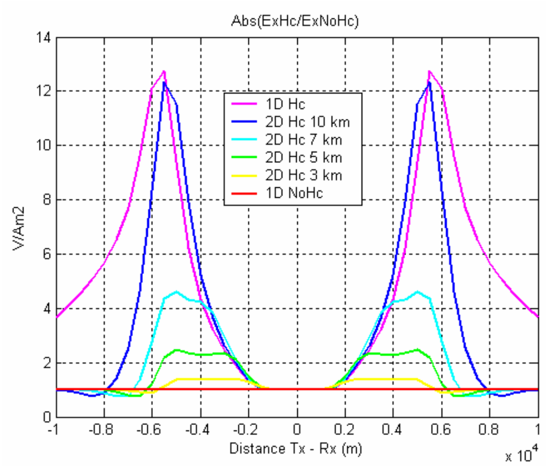

Figure 8 - Normalized in-line electrical field

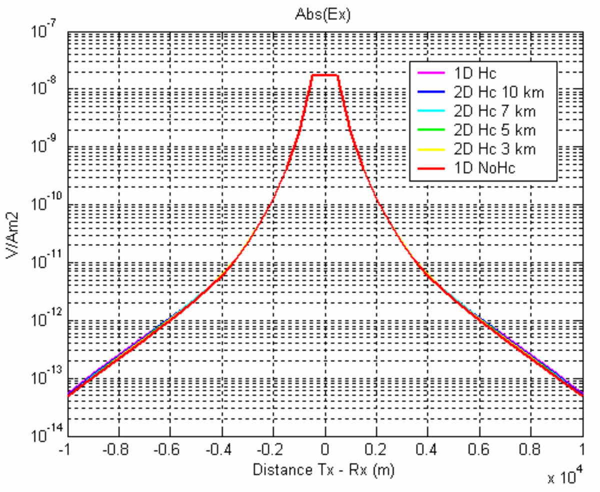

Figure 9 - Amplitude of the in-line electrical field

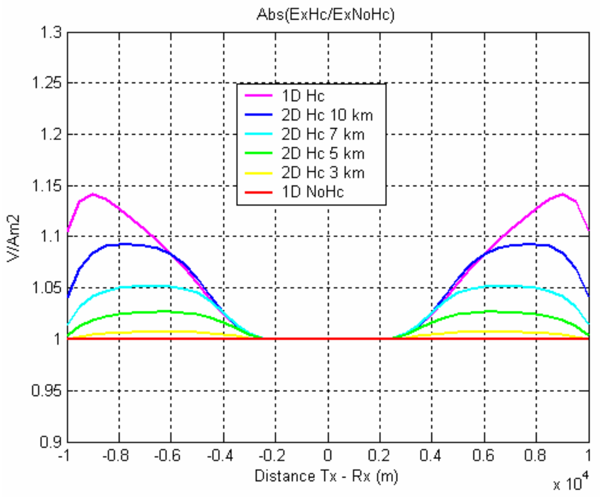

Figure 10 - Normalized in-line electrical field

Statistic of the experiments

The Figure 11 shows the performance in contrast with the Gaussian elimination and the preconditioned biconjugate gradient for the test model introduced in Figure 2. The bar graphics shows the computational time versus number of unknowns. At the first chart is demonstrated that Gaussian elimination solver is more efficient that the BCG algorithm, because in this experiment the band of matrix is constant, at the second graphic the band is not more constant and the BCG is more efficient, and finally if we have a huge quantity of unknowns the BCG is more recommended. The others test models had performed with equivalent results.

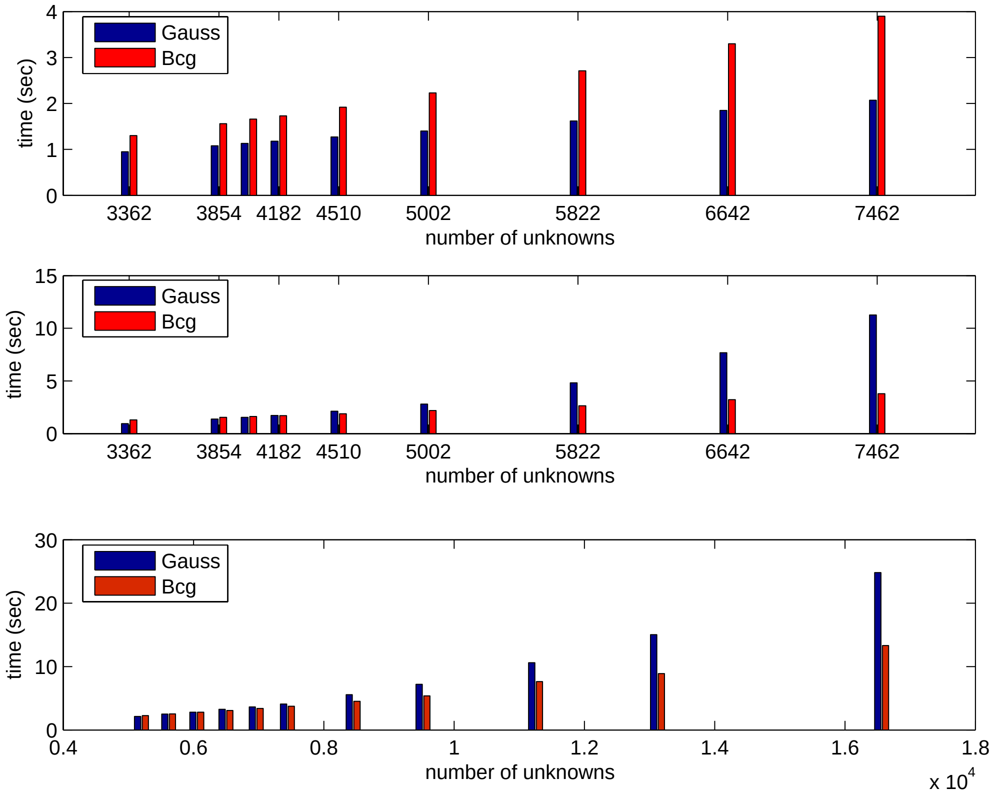

Figure 11 – The statistic of the performance of preconditioned biconjugate gradient in contrast with the Gaussian elimination.

Conclusions

The preconditioned biconjugate gradient method is suitable option as an iterative solver for electromagnetic modeling, it was demonstrated that in problems which the band of matrix is constant the Gauss algorithm is more efficient than the BCG method. On the other hand, in problems that have very large of the unknowns, the BCG method has a gain in performance in relation to the Gaussian elimination. Hence, with the strong tendency of electromagnetic problems in geophysics migrates to three-dimensional cases, where the number of unknowns is very large, the preconditioned biconjugate algorithm is a natural option to solve such problems.

References

Alumbaugh, D. L., Morrison, H. F., 1995. Theoretical and practical considerations for crosswell electromagnetic tomography assuming a cylindrical geometry: Geophysics, vol. 60, No. 3, p 846 -870.

Axelsson, O., Barker, V. A., 1984, Finite element solution of boundary value problems, Academic Press.

Axelsson, O., 2000. Iterative solution methods. Cambridge University Press.

Benzi, M.,2002. Preconditioning Techniques for Large Linear Systems: A Survey. Journal of Computational Physics, 182, p418- 477.

Coggon, J. H., 1971. Electromagnetic and electrical modelling by the finite element method. Geophysics 36, p132- 155.

Ellingsrud, S., Eidesmo, T., Johansen, S., Sinha, M.C., MacGregor, L.M. & Constable, S., 2002. Remote sensing of hydrocarbon layers using sea-bed logging (SBL): Results of a cruise offshore West Africa, The Leading Edge, 21, 972-982.

Eidesmo, T., Ellingsrud, S., MacGregor, L. M., Constable S., Sinha, M. C., Johansen S., Kong, F. N., and. Westerdahl, H., 2000. Seabed logging (SBL), a new method for remote and direct identification of hydrocarbon filled layers in deepwater areas, First Break, 20, p144- 152.

Freund, R. W., 1992. Conjugate gradient-type methods for linear systems with complex symmetric coefficient matrices. SIAM Journal on Scientific and Statistical Computing, 13, p425- 448.

Hestenes, **M. R. and E. Stiefel.**, 1952. Methods of conjugate gradient for solving linear systems. J. Res. Nat. Bur. Stand. 49, p409- 436.

Jacobs, D. A. H., 1981, The exploitation of sparsity by iterative methods, in Duff, I. S., Sparse matrices and their uses, academic Press, p191– 222.

Meijerink, J. A., and van der Vorst, H. A., 1977. An iterative solution method for linear systems for which the coefficient matrix is an M-matrix. Mathm. Comp. 31, p148- 162.

Pridmore, D. F., Hohmann, G. W., Ward S. H., and Sill, W. R., 1981, An investigation of finite – element modeling for electrical and electromagnetic data in three dimensions: Geophysics, vol. 46, No. 7, p 1009 -1024.

Rijo, L.,1977. Modeling of electric and electromagnetic data, Ph.D. Thesis. University of Utah.

Saad, Y., 1996. Iterative methods for linear systems. PWS Publishing Co.

Silva, M. W. C. and Rijo, L. 2003. Modeling the geo-magnetic fields under the Equatorial Electrojet perpendicular to the strike of a 2-D structure. In: 8th International Congress of the Brazilian Geophysical Society. Rio de Janeiro, SBGF. In CD-ROM.

Souza, V. C. T. and Rijo, L., 2003, The potentiality of the finite element method in electromagnetic tomography, 8th International Congress of The Brazilian Geophysical Society, Salvador, in CD-ROM

Smith, J. T., 1996. Conservative modeling of 3-D electromagnetic fields, Part II: Biconjugate gradient solution and an accelerator. Geohysics 61, p1319-1324.

Smith, C. F., Peterson, A. F., and Mittra, R., 1990. The biconjugate gradient method for electromagnetic scattering. IEEE Trans. Antennas Propagat., 38, p938- 940.

van der Vorst, H. A., 2000. Iterative method for large linear systems.

http://www.math.ruu.nl/people/vorst/lecture.htm

Wu, X., 2003. A 3-D finite-element algorithm for Dc resistivity modelling using the shifted incomplete Choleski conjugate-gradient method. Geophysical J. Int., 154, p947- 956.

Wu, X.; Xiao, Y.; Qi, C.,and Wang, T., 2003. Computations of secondary potential for 3D DC resistivity modelling using an incomplete Choleski conjugate-gradient method. Geophysical Prospecting, 51, p567- 577.

Zhang, J.,Mackie R. L. and Madden, T. R., 1995. Three-dimensional resistivity forward modeling and inversion using conjugate gradients. Geophysics 60, p1301– 1307.

Acknowledgments

This work was supported by Programa de Recursos Humanos da ANP - PRH/ANP-06 and CTPETRO/CNPq-46295/00-8. The authors would like to thank to Curso de Pós-Graduação em Geofísica – UFPA for logistic support and to all researchers this department.