NATIONAL INSTITUTE FOR SPACE RESEARCH – INPE/MCT
SOUTHERN REGIONAL SPACE RESEARCH CENTER – CRSPE/INPE - MCT
SOUTHERN SPACE OBSERVATORY – OES/CRSPE/INPE - MCT

FEDERAL UNIVERSITY OF SANTA MARIA – UFSM
SPACE SCIENCE LABORATORY OF SANTA MARIA – LACESM/CT - UFSM
INPE/MCT – UFSM/MEC

STATISTICAL STUDY OF CONSTRAINING PARAMETERS SPACE IN LEAST SQUARE METHOD APPLIED TO GAUSSIANS FIT TO POWER SPECTRA OF SIMULATED RADAR ECHOES OF ELECTROJET PLASMA IRREGULARITIES*

Henrique Carlotto Aveiro[1,2], Clezio Marcos Denardini[3], Mangalathayil Ali Abdu[3], Nelson Jorge Schuch[1]

(1) Ionospheric and Neutral Atmosphere Sounding Laboratory, LSIANT/CRSPE/INPE – MCT, Santa Maria - RS, Brazil;
(2) Space Science Laboratory of Santa Maria, LACESM/CT – UFSM, Santa Maria - RS, Brazil;
(3) Division of Aeronomy, DAE/CEA/INPE – MCT, São José dos Campos - SP, Brazil.
Contato: **aveiro@lacesm.ufsm.br** / Fax: +55 55 3220 8021

*COM2-01630-2005

ABSTRACT

The Least Square Fit Method modified by Levenberg-Marquardt is intensively used in many areas of space sciences for fittings mathematical equations to sample data. In equatorial electrojet studies based on back-scatter coherent radar data this method has been used to fit the sum of two Gaussians to the power spectra of the echoes back-scattered from 3-meter plasma irregularities. This fitting is done in order to identify the Doppler velocities of the two types of plasma irregularities power spectra present in the electrojet echoes. For the present work we have simulated coherent radar echoes to generate power spectra having the characteristic of the electrojet irregularities. The simulations were made based on *a priori* parameters of the plasma irregularities such as group velocity and spectral width. Using this simulated data set, we have performed a statistical study of constraining the parameters space in the Least Square Fit Method applied to two Gaussians fit to power spectra. The variance distribution of the Gaussian parameters is presented and analyzed in terms of the weights of the penalty functions, which prevent the method to converge to a local minimum. An analysis of the variance of each parameter as a function of its constraining value is also presented.

INTRODUCTION

At about 105 km of altitude in the equatorial E region and covering a latitudinal range of ±3º around the dip equator flows an intense electric current named equatorial electrojet (EEJ) driven by the E region dynamo [1, 2]. Studies of the equatorial ionosphere using VHF radars have shown echoes back-scattered from plasma irregularities in the EEJ which have shown distinct spectral signature for two observed irregularities, Type 1 and Type 2, also known as two-stream [3, 4] and gradient drift [5], respectively. They have been studied in order to explain the phenomenology [1, 2, 6, and references therein] and also in order to understand the E region electric fields [7, 8, 9]. Since 1998, when the Brazilian 50 MHz coherent back-scatter became fully operational, such studies have also been conducted in the Brazilian longitude sector [10, 11, 12, 13]. For such studies, a precise determination of the Doppler shift of the irregularities is a crucial requirement. And the curve fitting is presented as an efficient tool to determine the irregularity Doppler shift [14, 15, 16]. The curve fitting as a parameter estimation technique is based on finding the parameters of a well known mathematic equation, trying to minimize the mean square errors between observational data and the fit curve [17]. In this work we studied the implications of constraining the parameters of Gaussian curves fitted to power spectra of simulated back-scatter radar signals from Type 1 irregularities. This statistical study aims to quantify the advantages and disadvantages of applying such technique.

THEORY AND METHOD

This work focuses on the study of Type 1 power spectrum of the back-scattered signals from 3-m EEJ plasma irregularities that should present a sharp peak centered at around 120 Hz corresponding to the radar frequency of 50MHz. The Gaussian covariance model of Zrnic [18] was used to simulate power spectra of 3-m plasma irregularities containing both the characteristics of the Type 1 and Type 2 instabilities, each spectra having 256 points. Type 2 irregularities were simulated with Doppler frequency f_{d2} = 80Hz, standard deviation s_2= 50Hz and signal-to-noise ratio SNR_2 = 3dB. For Type 1 irregularities we have simulated three data set having the same Doppler frequency (f_{d1}) and standard deviation (s_1), namely, 120Hz e 20Hz, respectively. However, each set of Type 1 spectra was chosen to have different signal-to-noise ratio (SNR_1): 3dB, 6dB and 9dB. Thereafter, we have summed each spectrum of the three sets of Type 1 spectra with one spectrum of the set of Type 2 spectra. This resulted in three sets of simulated EEJ plasma irregularities spectra: 1) Type 2 with 3dB plus Type 1 with 3dB; 2) Type 2 with 3dB plus Type 1 with 6dB and 3) Type 2 with 3dB plus Type 1 with 9dB. White noise was added to signals in time domain in order to assure a more realistic variance in the power spectra. In this way, each simulated spectrum was described by the sum of two Gaussians and a noise level, i.e., it was described by one function S in relation to the frequency *f*, given by:

$$S(f) = \frac{P_1}{\sigma_1\sqrt{2\pi}} \exp\left[-\frac{(f - f_{d1})^2}{2\sigma_1^2}\right] + \frac{P_2}{\sigma_2\sqrt{2\pi}} \exp\left[-\frac{(f - f_{d2})^2}{2\sigma_2^2}\right] + P_N \qquad (1)$$

where P_1 and P_2 are, respectively, the areas of the Gaussians representing Type 1 and Type 2 spectra, P_N is the noise power, and the other parameters have been described above. To determine the 7 parameters of each spectrum, $\boldsymbol{a} = \{f_{d1}, f_{d2}, s_1, s_2, P_1, P_2, P_N\}$, the Maximum Likelihood Estimate (MLE) was used for nonlinear fitting. This method is based on finding the parameters $\boldsymbol{a}$ that maximize the probability function $P(y_1 \ldots y_n|\boldsymbol{a})$ of observing the data set $\boldsymbol{y} = \{y_1 \ldots y_n\}$. It is also a problem of finding the parameters $\boldsymbol{a}$ that minimize the square sum of residual errors between the data set $\boldsymbol{y}$ and the Gaussians S(*f*), considering the uncertainty s_i related to each point y_i. In view of this, equation (2) below presents our objective function to be minimized. Here *N* is the number of frequency points, y_i is the observed spectral amplitude for one given frequency in the power spectrum and all the other parameters have been introduced before.

$$\chi^2 = \sum_{i=1}^{N}\left[\frac{y_i - S(f_i; P_1, f_{d1}, \sigma_1, P_2, f_{d2}, \sigma_2, P_N)}{\sigma_i}\right]^2 \qquad (2)$$

Constraining the parameters search means to impose boundaries in the space of parameters, which can not be crossed by the method during the search. One way of doing this is by imposing penalties to the objective function when the method assumes unrealistic physical values [19]. We have done this by adding the following function to (2):

$$\Phi_j = \frac{\alpha_j}{h_j(P_1, f_{d1}, \sigma_1, P_2, f_{d2}, \sigma_2, P_N)} \qquad (3)$$

where a_j is the weight of the function and h_j is penalty function for each Gaussian parameter. Such function should be positive in the valid region of search, decrease rapidly as the search approach of the prohibited region, and be negative when the method crosses the boundary. In this way, the search for Gaussian parameters of spectra characteristics of irregularities Type 1 and Type 2 are limited to physically acceptable values. For this study we have used the method described above to simulated 3000 spectra characteristic of Type 1 and Type 2 EEJ irregularities, and separated them in three different groups (having 1000 spectra each one) according the SNR_1. Afterwards, every spectrum was fitted by two Gaussians, constraining *fd1* between 100 and 150 Hz, which corresponds to a Doppler velocity between 300 and 450 m/s. Moreover, we have studied the response of the fitting by constraining the parameter with five different weights in the penalty function: 10^0, 10^{-1}, 10^{-2}, 10^{-3} and 0. The basic analysis consisted in a direct comparison of the f_{d1} fitted to the spectra using Gaussians curves with the *a priori* f_{d1} value used to generate the Type 1 spectra. We have also compared the behavior of the variance of the f_{d1} and the standard deviation of the Gaussian Type 1 (sd_1) as the weight of the penalty functions (a_j) and power-to-noise level of the Gaussian Type 1 (SNR_1) increase.

ACKNOWLEDGEMENTS

The author wishes to thank CNPq/MCT for the support to his program through the project nr. 107616/2003-3. He also would like to acknowledge the Young Scientist Award offered by the International Union of Radio Science (URSI) providing the financial support to present the work.

RESULTS AND DISCUSSIONS

The distributions of the Doppler frequency of the Gaussian Type 1 (f_{d1}) as it was estimated by the MLE Method for different weights (a_j) in the penalty function is presented in Fig. 1. The results for the first data set, Fig. 1-a, shows when spectra Type 1 were simulated using SNR_1 = 3dB. Fig. 1-b shows the distribution for the case when SNR_1 = 6dB and Fig. 1-c gives the distribution of f_{d1} versus a_j when SNR_1 = 9dB. In this figure we see the distribution of *fd1* do not change much as a_j decreased from 10^0 to 10^{-3}. However, some differences can be noted in the distribution of f_{d1} when no penalty functions are applied. It seems the distribution spreads out from the expected frequency of 120Hz when applying constraints. This effect becomes notable when SNR_1 is comparable to SNR_2. We also see that as SNR_1 goes from 3 to 9dB the method clearly increased the percentage of spectra fitted with the right f_{d1}. This seems expected, but we should remember the increase in the signal power would increase the variance in the peak of the power spectrum, which in turn increases the uncertainties of the fitting method. Nevertheless, the MLE Method provides good results even when the data points have a large variance, which seems to be the case here.

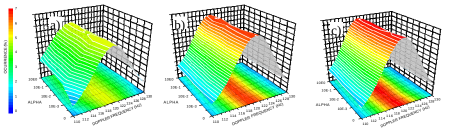

Fig. 1. Distribution of Doppler frequency type 1 versus *aj* for SNR1 equal to (a) 3dB, (b) 6dB and (c) 9dB.

The distributions of the variance of Doppler frequency estimated from the Gaussian Type 1 (Vf_{d1}) associated with the fitting method for different penalty function weights (a_j) is shown in Fig. 2. Like before, Fig. 2-a shows the results for SNR_1 = 3dB, Fig. 2-b shows the distribution for SNR_1 = 6dB, and Fig. 2-c gives the distribution for 9dB. In all these figures the low values of Vf_{d1} (<1) increase with applying constraints, which would lead to the interpretation that f_{d1} is better estimated. However, when applying constraints we have imposed penalties to the objective function by adding (3) with a_j in the numerator. Once Vf_{d1} is derived from the inverse of the gradient of the objective function, the higher a_j the lower Vf_{d1}. Thereafter, the above assumption based on Fig. 2 that f_{d1} is better estimated when using constraints should not be considering conclusive. Moreover, Fig. 1 shows that distribution of f_{d1} spreads out as a_j differs from zero. A tentative explanation for the increase shown in Fig. 2 is thought in terms of fitting sharp peaks.

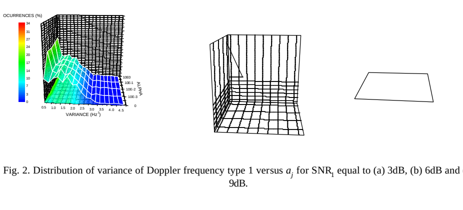

Fig. 2. Distribution of variance of Doppler frequency type 1 versus a_j for SNR_1 equal to (a) 3dB, (b) 6dB and (c) 9dB.

The distributions of the standard deviation (s_1) estimated for the Gaussian curve Type 1 for different penalty function weights (a_j) and SNR_1 are presented in Fig.3-a, Fig. 3-b and Fig. 3-c, respectively to 3dB, 6dB and 9dB. Here we see the distribution of the s_1 gets sharp as SNR_1 increases from 3 to 9dB. Each individual graph shows the occurrence of s_1 close to 20 Hz do not change much as a_j decreased from 10^0 to 10^{-3}. But some difference is observed in the distribution of s_1 without constraints. The distribution spreads in the base and decreases in amplitude. It indicates that the precision in estimating s_1 increases when a_j differs from zero.

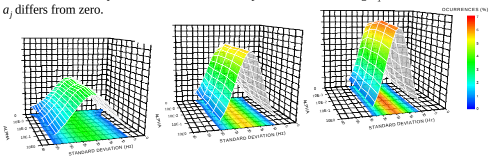

Fig. 3. Distribution of Standard Deviation type 1 in function of *aj* for SNR1 equal to (a) 3dB, (b) 6dB and (c) 9dB.

However, as shown in Fig. 4 below, the good precision in estimating s_1 is balanced by the reduction in the number of fitted spectra. This figure shows the percentage of fitted spectra in function of a_j for SNR_1 equal to (black) 3dB, (navy) 6dB and (blue) 9dB. Anyway, the increases in the precision in estimating s_1 seem to be the best result of applying constraints.

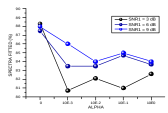

Fig. 4. Percentage of fitted spectra in function of *aj* for SNR1 equal to (black) 3dB, (navy) 6dB and (blue) 9dB.

CONCLUSIONS

We have studied the effect of constraining the range of Gaussian curves parameters during curve fitting methods applied to simulated power spectra having signatures of the presence of the well known Type 1 and Type 2 plasma irregularities from the equatorial electrojet. The main conclusion is that the application of constraints could compromise the estimation of the Doppler frequency f_{d1}, increasing the uncertainty around the "right" value. On the other hand, the standard deviation of the Gaussian Type 1 seems to be better fitted, as was presented in the Fig. 2. The results have shown the application of constraints apparently increases the variance of the curve, because the distributions of frequency had spread out when a_j differed from zero. Maybe the results can be consequences of using a non-adaptable a. Other alternative is increase the maximum number of iterations before accept the non-convergence. Despite these conclusions seem to be simple, they have direct impact into the data analysis from coherent back-scatter radars echoes. An application of constraint without consider what we have present here would not compromise seriously any result based on the Doppler velocity of Type 1 irregularities. But it will certainly lead to a higher error and maybe to error propagations into the determination of dependent quantities like electric field inferred from the Doppler velocities from such irregularities. Finally, we should mention that the choice for constraining the parameters should be related to the objective of the research and its can not be generalized for all the cases.

REFERENCES

[01] **Forbes, J. M.**, 1981, The Equatorial Electrojet: Reviews of Geophysics and Space Physics, Vol. 19, No. 3, p.469-504.
[02] **Fejer, B. G. and Kelley, M. C.**, 1980, Ionospheric Irregularities: Reviews of Geophysics and Space Physics, Vol. 18, No. 2, p.401-454.
[03] **Farley, D. T.**, 1963, A plasma instability resulting in field aligned irregularities in the ionosphere: Journal of Geophysical Research, Vol. 68, No. A22, p.6083-6097.
[04] **Buneman, O.**, 1963. Excitation of field aligned sound waves by electron streams: Physical Review Letters, Vol. 10, No. 7, p.285-287.
[05] **Rogister, A. and N. D'Angelo**, Type II irregularities in the equatorial electrojet, J. Geophys. Res., 75, 3879-3887, 1970.
[06] **Reddy, C. A.**, 1981, The equatorial electrojet: a review of ionospheric and geomagnetic aspects: Journal of Atmospheric and Terrestrial Physics, Vol. 43, No. 5/6, p.557-571.
[07] **Reddy, C. A., Vikramkumar, B. T. and Viswanathan, K. S.**, 1987, Electric fields and currents in the equatorial electrojet deduced from VHF radar observations – I. A method of estimating electric fields: Journal of Atmospheric and Terrestrial Physics, Vol. 49, No. 2, p.183-191.
[08] **Hysell, D. L., Larsen, M. F. and Woodman, R. F.**, 1997, JULIA radar studies of the electric field in the equatorial electrojet: Geophysical Research Letters, Vol. 24, No. 13, p.1687-1690.
[09] **Hysell, D. L. and Burcham, J. D.**, 2000, Ionospheric electric field estimates from radar observations of the equatorial electrojet: Journal of Geophysical Research, Vol. 105, No. A2, p.2443-2460.
[10] **Abdu, M. A., Denardini, C. M., Sobral, J. H. A., Batista, I. S., Muralikrishna, P. and de Paula, E. R.**, 2002, Equatorial electrojet irregularities investigations using a 50 MHz back-scatter radar and a Digisonde at São Luís: Some initial results: Journal of Atmospheric and Solar-Terrestrial Physics, Vol. 64, No. 12-14, p.1425-1434.
[11] **Abdu, M. A., Denardini, C. M., Sobral, J. H. A., Batista, I. S., Muralikrishna P., Iyer, K. N., Veliz, O. and de Paula, E. R.**, 2003, Equatorial electrojet 3 m irregularity dynamics during magnetic disturbances over Brazil: results from the new VHF radar at São Luís: Journal of Atmospheric and Solar-Terrestrial Physics, Vol. 65, No. 14-15, p.1293-1308.
[12] **Denardini, C. M., Abdu, M. A. and Sobral, J. H. A.**, 2004, VHF radar studies of the equatorial electrojet 3-m irregularities over São Luís: day-to-day variabilities under auroral activity and quiet conditions: Journal of Atmospheric and Solar-Terrestrial Physics, Vol. 66, n17, p.1603-1613.
[13] **Denardini, C. M.**, 2004, Estudo da eletrodinâmica da ionosfera equatorial durante o período de máxima atividade solar (1999-2002): São José dos Campos, Instituto Nacional de Pesquisas Espaciais, Thesis.
[14] **Press, W. H., Teukolsky, S. A., Vetterling, W. T., Flannery, B. P.** Numerical recipes in C: the art of scientific computing. 2. ed. Cambridge: Cambridge University Press, 1992.
[15] **Kudeki, E., Bhattacharyya, S. and Woodman, R. F.**, 1999, A new approach in incoherent scatter F region E x B drift measurements at Jicamarca: Journal of Geophysical Research, Vol. 104, No. A12, p.28145-28162.
[16] **Denardini, C. M.**, 2003, Estimação de parâmetros de dados físicos: São José dos Campos, Instituto Nacional de Pesquisas Espaciais.
[17] **Bevington, P. R. and Robinson, D. K.**, 1992, Data reduction and error analysis for the physical science. McGraw-Hill.
[18] **Zrnic, D. S.**, 1979, Estimation of spectral moments of weather echoes: IEEE Transactions of Geosciences on Electronics, Vol. 17, p.113-128.
[19] **Bard, Y.**, 1974, Non linear parameter estimation. Academic Press.