URSI AP-RASC 2019, New Delhi, India, 09 - 15 March 2019

An Automatic Ionogram Scaling Technique based on the Nelder-Mead method

Konstantin Tsybulya, Nadezhda Kotonaeva, Anna Polyakova
Fiodorov Institute of Applied Geophysics, Moscow, Russian Federation, http://www.geospace.ru

Abstract

The work describes an approach to automatic ionogram scaling, suitable for both bottomside and topside sounding. The algorithm is designed to be used on high-perfomance parallel computing servers. It employs ionogram modeling by simulating radio wave propagation in plasma layers and adjusting the parameters of the layers by Nelder–Mead algorithm until the optimal coincidence with the results of sounding is achieved. The validation of automatic scaling was carried out for F2 layer critical frequency (f_oF2) by comparison with the data sequence for all previous observation years for a given station. The f_oF2 deviation quantiles were calculated for different geomagnetic activity levels. An ionogram's automatic scaling was considered a failure if the result fell in the lowest 0.01 or the highest 0.99 quantile.

1. Introduction

The most common approaches to create algorithms for ionogram-scaling software assume that this software runs at the sounding site. However, today it is possible to use high-performance servers, especially those designed for parallel computing. If an ionogram sounding network has a central processing and archival site, it is possible to install a powerful parallel computational server there and run the scaling software on it. Besides, there are plans to launch a constellation of ionosphere-sounding satellites that would amend Roshydromet ground-based network with topside-sounding data. Such space-based system would produce an ionogram every few seconds. To make it practically useful for space weather analysis and forecast, it is necessary to create an automatic scaling system, able to reduce this stream of data to the numerical characteristics of the ionosphere layers.

Thus, there is a need in automatic scaling software that is able to utilize modern high-performance parallel computation servers and could be easily adapted to scale the topside sounding ionograms. To satisfy this need, we propose a technique of automatic ionogram scaling based upon, first, modeling the ionograms by simulating the wave propagation and, second, fitting the modeled ionograms to the real ones by solving the optimization problem with the help of the Nelder–Mead method.

2. A short description of the algorithm

For most practical purposes, the ionosphere is considered to consist of a number of flat layers without strong horizontal gradients. In fact, the space weather situational reports supply a set of standard ionospheric parameters like the layers' critical frequencies, peak heights and some measure of the layer thickness [1]. Therefore, the algorithm begins with a set of standard ionospheric parameters, describing the layers. Based upon this set, one can construct a vertical profile of the ionospheric plasma density. One assumes that individual layers have some simple form, for example parabolic or sinusoidal. Now, having a profile, it is possible to simulate step-by-step the propagation of a radio wave pulses on all ionosonde frequency channels, computing the radar delay for each of them and thus constructing a simulated ionogram. This ionogram, with the trace width set to the sounder pulse length, is used as a template superimposed on the analyzed ionogram. At this point one can estimate how well the simulated and real ionograms coincide, that is, what fraction of the sounder-obtained ionogram falls in the modeled template.

By changing the values of the initial profile-describing parameters, one increases or decreases the degree of coincidence. Finally, one may find the template having maximal coincidence with the real ionogram. The ionospheric parameters, which were used to construct this optimal template, are considered the best description of the ionosphere over the sounding site.

Speaking mathematically [2], the scaling task is a solution of an optimization problem: to find the global maximum of a function in a multi-dimensional space. The ionospheric parameters are arguments of the function (and dimensions of the space). The metric that describes the degree of coincidence between the real and modeled is the function's result.

To solve the problem we decided to use the Nelder–Mead algorithm [3]. The main advantage of this method is that there is no necessity to calculate derivatives. In addition, this algorithm is easily adapted for parallel computing, since several ionograms may be computed simultaneously, as well as the pulse propagation on all frequency channels may be simulated in parallel.

From other side, this algorithm has tendency to converge to false (local) maximums. In these maximums, any variation of the ionospheric parameters makes the coincidence worse, though the template is very far away from the real ionogram traces.

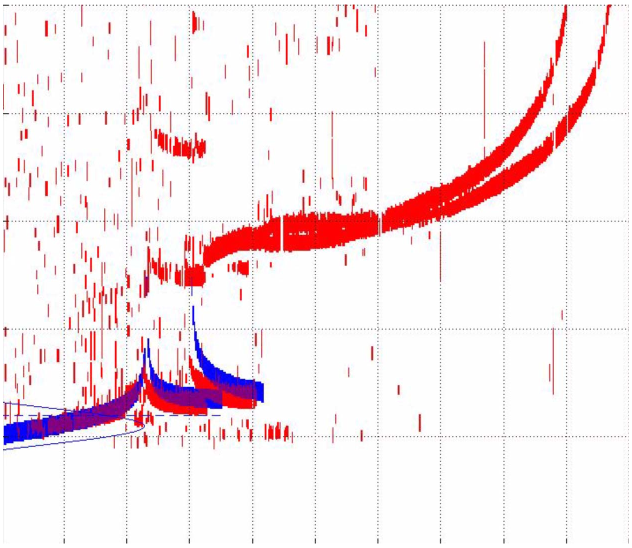

Figure 1. Early stage of the E-region template fitting. Blue templates partially cover red E-layer O- and X-traces. Also E-layer profile (solid line) and sporadic E position (dashed line) are shown.

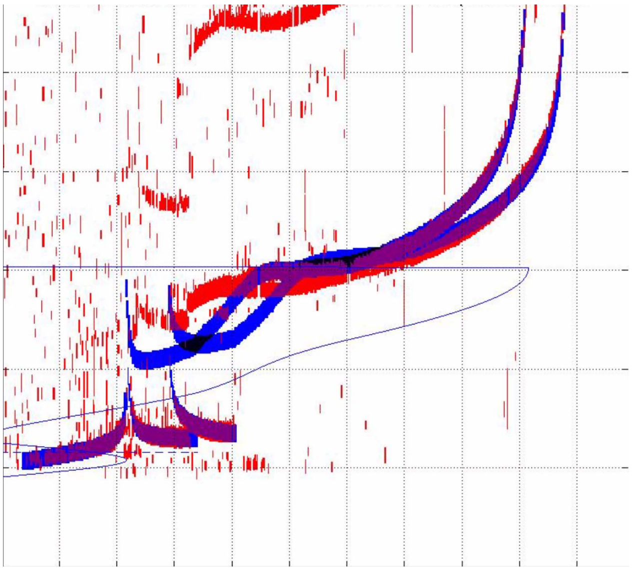

Figure 2. Intermediate stage of the F-region template fitting. The optimal E-region configuration is already found. The profile is shown only below the main maximum.

To overcome this problem we, again, use the advantage of parallel computing. The range of each parameter is divided on several subranges, thus all the parameter space is filled with a regular grid. Each cell of the grid is checked separately on the presence of maximums. Having more processing power means more reliable recognition of the ionogram since the grid may be made denser.

3. Some technical details

In some cases, the template runs very close to an ionogram trace yet does not intersect it. The coincidence metric is formally zero, since no radar response lies inside the template. To avoid the problem, the sharp borders of the trace are “blurred” by adding some low intensity wings, so the comparison metric grows if the template is brought closer to the trace.

Three parameters for every layer (F2, F1 and E) plus some additional characteristics of sporadic E layer and spread F features bring the total number of dimensions of the space to 12 or more. Finding maximums in such multidimensional space presents a challenge, so it is better to optimize the algorithm. For example, one can search for the best coincidence separately in the E region (Figure 1) and in the high frequency asymptotic tails zone. Then it is possible to connect the two zones by finding the best fit of the F1 layer (Figure 2).

An initial estimation of the parameter ranges are taken from the ionospheric models (for example, IRI). Also some optimization is possible at this stage: there is no need to look for the E and F1 layers at night time. The technique is very flexible; it allows changing the number of the ionosphere-describing parameters. Additional parameters allow better template fit but require higher performance computers.

Yet no mathematical function can describe real ionospheric profiles. So even in the case of smooth, horizontally layered ionosphere there always will be some deviation of the ionogram traces from the modeled templates. If there is a need of real-height ionospheric profiles, derived from the ionogram, then there must be an additional correcting stage to move the virtual height points up or down to the lower edge of the trace.

Finally, it is rather easy to adapt the method for the satellite sounding. In fact, the topside version is even simpler, since the complicated features of the lower ionosphere are not seen on the topside ionograms.

4. Algorithm validation

To check the validity of the automatic scaling algorithm and find errors, including those caused by spread F, travelling ionospheric disturbances, inclined layers, radio interference etc., we calculated the quantiles of measured f_oF2 deviations from the 27-day moving median.

As a measure of geomagnetic activity, we chose the index $a_p(t,\tau)$, which takes into account the recent history of the standard $a_p(t)$ index:

$$a_p(t,\tau) = (1-\tau)\cdot \sum_{n=0}^{80} a_p(t-3n)\cdot \tau^n \ , \qquad (1)$$

where $\tau = \exp(-3/\tau_0)$, and $a_p(t)$ – the standard 3-hours geomagnetic index in the time moment t;

τ_0 = 10 –characteristic time of the thermospheric reaction on geomagnetic activity surges, in hours.

The f_oF2 samples were taken for a given local time hour and month, only for quiet geomagnetic conditions with $a_p(t,\tau) < 10$, for which the greatest volume of data is available.

The distribution of the deviations was found to be Gaussian or close to it. The histograms for local midnight and noon deviations for the Moscow station, month March, are shown on Figure 3 and 4.

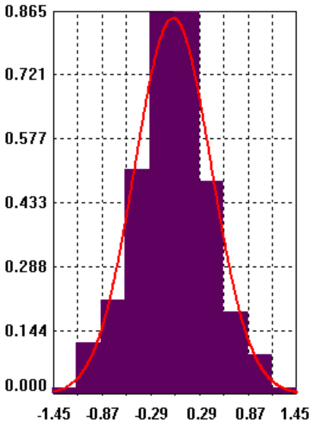

Figure 3. The deviation histogram of automatically derived f_oF2 values (in MHz). Quiet geomagnetic conditions, month March, local midnight.

The Pearson's chi-squared test shows that for the midnight case the distribution of the f_oF2 deviations from median is Gaussian with the mean of –0.02 MHz and the variance of 0.44 MHz.

For the local noon case, the distribution is close to Gaussian, with the mean 0.02 MHz and variance 0.53 MHz.

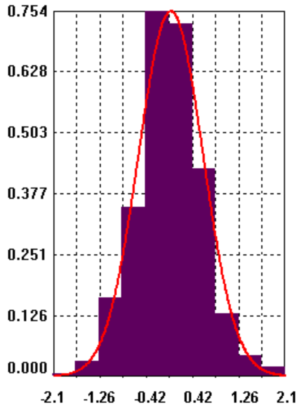

Figure 4. The deviation histogram of automatically derived f_oF2 values (in MHz). Quiet geomagnetic conditions, month March, local noon

Statistics shows that deviation quantiles from the level of 0.5 to the level of 0.95 strongly depend on local time and season. For some mid-latitude stations with long observation history, as Moscow, these diurnal and seasonal variations are seen even for the quantiles of higher levels (0.99 and 0.01).

The diurnal variation of f_oF2 deviation quantiles for the Moscow station, month March, under quiet geomagnetic conditions is represented in Figure 3 [4].

The solid lines 1 encompass all deviations from the median for the station for all its history, including operator errors and data transmission errors. The dashed lines 2 from above and below denote the 0.99 and 0.01 quantiles. The moment of 4 hours UT corresponds to the sunrise, so there is no negative deviations.

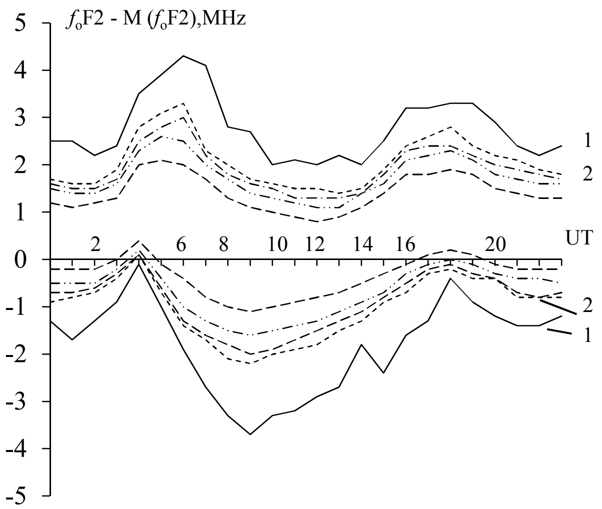

Figure 5. The diurnal variation of the f_oF2 deviation quantiles for quiet geomagnetic conditions, March, the station Moscow.

The automatic scaling validation procedure was based on the following rule: if the result of scaling falls behind the lines 1 and 2, then, with 98% probability, the ionogram is processed incorrectly and has to be checked and corrected manually.

The implementation of this control technique has shown for year 2018 that 1.9% need manual correction after the automatic scaling.

5. Conclusions

The suggested technique of automatic ionogram scaling allows utilization of the high-performance parallel computers power. The algorithm has to be attuned to the data; its flexibility allows doing it in many ways. The approach allows flexible changes of the algorithm to increase the range of recognized ionospheric features depending on processor power available.

The percentage of failures of the algorithm under quiet geomagnetic conditions for mid-latitude stations is around 1.9%.

6. Aknoledgements

The reported study was partially funded by RFBR grant for the research project № 18-05-80023

7. References

1. W.R. Piggot, and K.U. Rawer. URSI “Handbook of Ionogram Interpretation and Reduction,” UAG-23A, WDC-A for STP, Boulder, 1972.

2. R. O. Duda, P. E. Hart, and D.G. Stork. “Pattern Classification”. John Wiley & Sons, second edition. 1999.

3. J. A. Nelder, and R. A. Mead. “Simplex method for Function Minimization,” *Computer Journal*, **7**, 1965, pp. 308-313.

4. N. Kotonaeva, V. Mikhailov, E. Perminova “Statistic regularity of ionosphere critical frequency variation at quite geophysical conditions” XXV Russian conference. “Radio waves propagation” Tomsk. 2016. pp. 121-124.