A method for automatic detection of equatorial spread-F in ionograms

Carlo Scotto, Alessandro Ippolito, Dario Sabbagh

Istituto Nazionale di Geofisica e Vulcanologia, Rome, Italy (carlo.scotto@ingv.it)

A method is presented for automatic detection of spread-F. The method is based on an image recognition technique and is applied to ionograms recorded at the ionospheric station of Tucumán (-26.9°, 294.6°). The performance achieved is statistically evaluated and demonstrated with significant examples. The proposed method improves Autoscala's ability to reject ionograms with insufficient information, including those featuring Spread-F. Automatic identification of cases of spread-F is of additional interest in Space Weather applications, when it helps detect degraded radio propagation conditions.
The present data analysis is a retrospective study but forms the basis for real-time application as an extension of Autoscala's capabilities.

1. Introduction

The Italian National Institute of Geophysics and Volcanology (INGV) team have been working on the automatic interpretation of ionograms for several years and they have developed a computer program called Autoscala, which is capable of establishing the main ionospheric parameters. It is based on image processing techniques and is structured into several subroutines. Some subroutines identify the main traces: F2 (Scotto and Pezzopane, 2002; Pezzopane and Scotto, 2007), F1 (Pezzopane and Scotto, 2008), and Es (Scotto and Pezzopane, 2007). Other routines tackle specific problems, such as the elimination of double reflections (Scotto and Pezzopane, 2008), and highlighting the F2 trace (Pezzopane and Scotto, 2010). The vertical electron density profile is also estimated and returned as an output in a different procedure (Scotto, 2009).
The greatest difficulties in interpreting ionograms arise when the ionosphere is not ~~horizontally~~ vertically stratified. The ionosphere can be tilted so that reflections from several directions are possible at the same time. Field aligned irregularities can also give strong reflections and add to the complexity of the ionogram. These phenomena are particularly important at high and low magnetic latitudes and they give rise to undefined traces that make it difficult to recognize the critical frequency f_oF2.
According to the URSI standard, when the estimated uncertainty of a value is less than 2%, the numerical value is ~~unqualified~~ provided without qualificative letter. (Piggott and Rawer, 1972). Different types of spread F are classified according to Penndorf (1962). This classification is probably not complete and some of Penndorf's classes are not easily distinguishable (Piggott and Rawer, 1972). Nowadays, the range spread-F (RSF), strong range spread-F (SSF), frequency spread-F (FSF), and mixed spread-F (MSF) categories are more commonly used to describe F spreading (e.g. Shi et al., 2011), according to its appearance in the ionograms. The work presented in this paper addresses the problem of detecting the above-mentioned cases when, according to the URSI standard, it is impossible to assign a~~n unqualified~~ value for f_oF2 without qualificative letter, regardless of the spread F type.

Currently there are no programs capable of automatically identifying spread F in ionograms. Some attempts to automate identification have been made in polar ionograms (Scotto and Pezzopane, 2011), but it would be very useful to develop a routine to specifically identify equatorial spread-F. This is because day-to-day and short-term variability in Equatorial Spread-F (ESF), related to the development of plasma bubbles, is a major concern for space applications and operational systems (Abdu and Kherani, 2011). The correspondence between Strong Range Spread-F (SSF) and scintillations was also found to be systematic in an analysis of data recorded from October 2010 to September 2011 by an Advanced Ionospheric Sounder-INGV ionosonde, ~~and~~ a GPS Ionospheric

Scintillation, and total electron content (TEC) monitor scintillation receiver, located at Tucumán (-26.9°, 294.6°) (Alfonsi et al., 2013).

ESF is due to irregularities associated with Collisional Rayleigh Taylor (CRT) instability, which acts on the night time magnetic equatorial and low latitude F region (Woodman and Lahoz, 1976; Farley et al., 1970; Hysell and Burcham, 2002; Haerendel, 1973). Various studies demonstrated that the characteristics of ESF depend on longitude (Abdu, 2012, Ram et al., 2006; Manju et al., 2009; Manju et al., 2011; Wan et al., 2018). Chandra and Rastogi (1970) considered the existence of an equatorial belt of about ±20° geomagnetic latitude of high occurrence of Spread-F. They showed that, in this belt, ESF in the American and African sectors peaks during the local summer, while the minimum levels occurs during the local winter, independently of solar activity. Recently (Haridas et. al 2018) showed that the percentage occurrence and duration of ESF vary, respectively exponentially and linearly, with the Rayleigh Taylor (RT) instability growth rates over the magnetic equatorial location of Trivandrum (8.5° N, 77.0° E).

This paper describes a routine for identifying spread-F, applied to ionograms recorded by the Advanced Ionospheric Sounder, which was developed at the INGV and is installed at the ionospheric station of Tucumán.

2. Method

The method applied in this work is based on calculation of contrast, which is the same approach used by Autoscala for assessing the similarity of an empirical curve with a recorded trace. Initially an ionogram is loaded by Autoscala as a matrix M of N_x columns and N_y rows. N_x and N_y are defined by the following formulas:

$$N_x = \text{int}[(f_f - f_0)/\Delta f]+1$$

and

$$N_y = \text{int}[(h'_f - h'_0)/\Delta h']+1$$

where f_0, f_f, are respectively the ionosonde's lowest and highest frequency scanning range, Δf is the frequency step, h'_0, h'_f are respectively the lowest and highest virtual heights recorded on the ionogram. $\Delta h'$ is the height resolution at which the sounding was recorded. The entry $m_{i,j}$ (with i =1, . . ., N_x and j=1, . . ., N_y) of the matrix M is an integer proportional to the echo amplitude received by the ionosonde. This value is retrieved directly from the binary file recorded by the ionosonde.
Once the ionogram is memorized as a matrix of elements $m_{i,j}$, an empirical curve T capable of fitting the typical shape of the F2 trace is defined by the following parametric form:

$$T = \begin{cases} f = a - k \\ h' = \text{int}\left\{ H + A\tan\left[\dfrac{\pi}{2} \dfrac{\Delta x - k}{\Delta x} \right] \right\} \\ 0 \le k \le \Delta x \end{cases} \qquad (1)$$

In (1) the frequencies f and the virtual heights of h' are expressed as integers and correspond to the indices i and j of the matrix M. The parameters defining the curve are: H, a, and A. H is an integer varying from 1 to (N_y-30) and corresponds to the values of the horizontal asymptote for T. a is an integer varying from 1 to N_x, and corresponds to the value of the vertical asymptote for T. A is a decimal coefficient. Δx is an integer set to 20 representing the frequency interval, expressed in

pixels, where T develops. The curve starts from frequencies (a-Δx) and extends up to a. k is an integer varying from 0 to Δx. For small values of a, it can happen that f is negative: these values are intercepted and ignored as they represent negative frequency values. By varying a and H the curves are slid across the ionogram, while varying a and A generates different shapes for matching with the F2 trace.
For each curve T the local contrast $C(H, a, A, \Delta x)$ with the recorded ionogram is computed as:

$$C(H,a,A,\Delta x) = \sum_{k=a-\Delta x}^{k=a} SEC\left[k, T(H,A,\Delta x,k)\right] \quad (2)$$

where *SEC* is the Single Element of Contrast, which is described in Cesaroni et al., (2013). It is worth specifying that the curve T described here is not used by Autoscala to identify the F2 trace from an ionogram. This issue is addressed using a more sophisticated method specifically designed for the task (Scotto and Pezzopane, 2002; Pezzopane and Scotto, 2007).

For each ionogram, two parameters are taken into account:

1) C_{max}, the maximum value of C.

2) $C_{average}$, the average value of C.

If $C_{max}/ C_{average} \geq R_{threshold}$ then the algorithm assumes that the ionogram exhibits a spread-F phenomenon, while if $C_{max}/ C_{average} < R_{threshold}$ it assumes that it does not.

The problem therefore lies in choosing an appropriate $R_{threshold}$. This problem is dealt with in the next section with the application of the Receiver Operating Characteristic (ROC) curve method.

3. The ROC method

In signal detection theory, a ROC curve, as a binary classifier system, is a graphical plot of true positive rate (TPR) vs. false positive rate (FPR), as the discrimination threshold varies.
Each point on the ROC curve represents a TPR/FPR pair corresponding to a particular decision threshold. A test with perfect discrimination would have a ROC curve passing through the upper left corner (100% TPR, 0% FPR) of the plot. Therefore, the closer the ROC plot is to the upper left corner, the higher the overall accuracy of the test.
The present study considered a dataset A of 198 ionograms recorded by the AIS ionosonde installed at Tucumán, representing a wide range of situations. Within this dataset a subset F of ionograms was defined in which an operator had indicated the presence of spread-F, and a subset NF of ionograms was defined not affected by this phenomenon. For each possible $R_{threshold}$ value, a fraction of ionograms belonging to F were correctly classified by the filter as positive (TP = true positive), while the remainder of F were incorrectly classified by the filter as negative (FN = false negative). Similarly, for each possible $R_{threshold}$ value, a fraction of ionograms belonging to NF were correctly classified by the filter as negative (TN = true negative), while the remainder of NF were incorrectly classified as positive (FP= false positive). The ROC curve can quantify, in terms of TPR= TP/(TP+FN) and FPR = FP/(FP+TN), the overall effectiveness R of correctly discriminating the two F and NF ionogram populations.
Fig. 1 shows the ROC curve for the binary classifier described in this work. Based on this curve the $R_{threshold}$ was definitively set to 1250.

4. Testing the method

A test was performed using a set I of 7649 ionograms not included in the database used to select $R_{threshold}$, and corresponding to all available hourly ionograms recorded at Tucumán in 2016. These ionograms were manually scaled by an operator, who selected a subset of ionograms with spread-F and a subset without spread-F. The results obtained are summarized in Table 1, where the TP, FN, TN, and FP obtained for the set I are reported after applying the routine with the decisional threshold $R_{threshold}$ set to 1250. It emerged that the routine successfully detected a high percentage [TP/(TP+FN) = 81.55%] of spread-F cases. A small percentage of ionograms without spread-F were wrongly tagged as being affected by spread-F [FP/(TN+FP) = 0.55%].
Typical cases are reported in Fig. 2 and Fig. 3. Fig. 2 shows an ionogram without spread-F features correctly tagged (TN), and Fig. 3 a case of correctly identified equatorial spread-F (TP).
Fig. 4 and Fig. 5 report critical cases. Fig. 4 shows an ionogram in which the routine fails to detect a spread-F trace, misleading Autoscala. In this case an underestimated *foF2* value is given as output (FN case). Fig 5 presents an ionogram showing well defined traces. This ionogram is wrongly tagged as featuring spread-F and discarded (FP case). In this case no *foF2* value is provided as output by Autoscala.

5. Rejecting ionograms with insufficient information

A further test was performed using the same set I of 7649 ionograms outside the database used to select $R_{threshold}$. These ionograms were analyzed by an operator, who selected a subset R of ionograms with insufficient information to scale an *foF2* value, and a subset NR of ionograms for which *foF2* could be scaled. Autoscala already has an algorithm for the identification of ionograms with insufficient information, but these are rejected without providing any information on the various reasons of rejection, such as sporadic E-layer, spread-F, radio interference, or failure of the apparatus. The existing capability of Autoscala to reject ionograms with insufficient information was verified using the subsets R and NR as a reference, and the results are summarized in Table 2.
Autoscala's code was subsequently modified, implementing the initial routine developed in this work to identify cases of spread-F. Ionograms with spread-F can thus be preliminarily tagged and rejected. The algorithm for identifying ionograms with insufficient information can then be successfully applied to the ionograms in which spread-F is not detected, as in the previous version of Autoscala. The capability of the modified version of Autoscala to reject ionograms for which it is not possible to establish an *foF2* value - either specifically because of spread-F or generically due to insufficient information - was also verified using the subsets R and NR as a reference, and the results are summarized in Table 3.

Comparing the results presented in Table 2 and Table 3, it is possible to see that the original version of Autoscala correctly rejected 56.7% of ionograms, while the improved version correctly rejected 85.9%. The percentage of ionograms rejected by Autoscala but scaled by the operator are 2.0% for the original version of Autoscala and 2.3% for the modified version.

6. Summary

The routine described in this paper was able to identify ionograms from the Tucumán ionospheric station with equatorial spread-F: 81.55% of spread-F cases were successfully detected, and 0.55% of the ionograms without spread-F were wrongly tagged.

The routine can be applied to ionograms from the Tucumán ionospheric station to identify cases of spread-F and obtain a better overall rejection rate for ionograms for which an *foF2* value cannot be established as output. By implementing the routine described in this work into the Autoscala

program, it was demonstrated that the correct rejection percentage rises from 56.7% to 85.9%, while the incorrect rejection percentage rises only from 2.0% to 2.3%.

The current routine for spread-F identification is not able to separate ionograms showing range spread-F (RSF), strong range spread-F (SSF), frequency spread-F (FSF), and mixed spread-F (MSF). Achieving the above would open up interesting prospects for real-time assessment of the accuracy and reliability of navigation systems.

7. References

Penndorf, R., 1962. Classification of spread F ionograms, J. Atmos. Terr. Phys., 24 (9), 771-778.

Chandra, H., Rastogi, R., G., 1970. Solar cycle and seasonal variation of spread F near the magnetic equator, J. Atmos. Terr. Phys., 32, 439-443.

Farley, D.T., Balsley, B.B., Woodman, R.F., McClure, J.P., 1970. Equatorial spread F: implications of VHF radar observations. J. Geophys. Res. 75, 7199–7216.

Piggott, W., R., Rawer, K., 1972. URSI handbook of ionogram interpretation and reduction. World Data Center A for Solar-Terrestrial Physics, Report UAG - 23.

Haerendel, G., 1973. Theory of Equatorial Spread F, Report. Max-Plank Inst. Fur Phys. and Astrophys., Munich.

Hysell, D.L., Burcham, J., 2002. Long term studies of equatorial spread F using the JULIA radar at Jicamarca. J. Atmos. Terr. Phys. 64, 1531– 1543.

Scotto, C., Pezzopane, M., 2002. A software for automatic scaling of foF2 and MUF (3000) F2 from ionograms. URSI XXVIIth General Assembly, www.ursi.org/proceedings/procGA02/papers/p1018.pdf.

Woodman, R.F., Lahoz, C., 1976. Radar observations of F region equatorial irregularities. J. Geophys. Res. 81, 5447.

Ram S., T, , Rao, P.,V.,S. R., Niranjan, K., Prasad, D., S.,V.,V., D., Sridharan, R., Devasia, C., V., Ravindran, S., 2006. The role of post sunset vertical drifts at the equator in predicting the onset of VHF scintillations during high and low sunspot activity years. Ann. Geophys., 24, 1609-1616.

Pezzopane, M., Scotto, C., 2007. Automatic scaling of critical frequency foF2 and MUF (3000) F2: A comparison between Autoscala and ARTIST 4.5 on Rome data. Radio Sci. 42 (4), doi:10.1029/2006RS003581.

Scotto, C., Pezzopane M., 2007. A method for automatic scaling of sporadic E layers from ionograms. Radio Sci. 42 (2),. doi:10.1029/2006RS003461.

Pezzopane, M., Scotto, C.,-2008. A method for automatic scaling of F1 critical frequencies from ionograms. Radio Sci. 43 (2), doi:10.1029/2007RS003723.

Scotto, C., Pezzopane, M., 2008. Removing multiple reflections from the F2 layer to improve Autoscala performance. J. Atmos. Solar-Terr. Phys. 70 (15), 1929-1934. doi: 10.1016/j.jastp.2008.05.012.

Manju, C., V., Devasia, S., R., 2009. The seasonal and solar cycle variations of electron density gradient scale length, vertical drift and layer height during magnetically quiet days: implications for spread F over Trivandrum, India, Ann. Geophys., 27 503-510.

Scotto, C., 2009. Electron density profile calculation technique for Autoscala ionogram analysis. Adv. Space Res. 44 (6), 756-766. doi: 10.1016/j.asr.2009.04.037.

M., Pezzopane, Scotto, C., 2010. Highlighting the F2 trace on an ionogram to improve Autoscala performance. Computers & Geosciences. 36 (9), 1168-1177. doi: 10.1016/j.cageo.2010.01.010.

Manju, G., Sreeja, V., Ravindran, S., Thampi, S.,V., 2011. Toward prediction of L band scintillations in the equatorial ionization anomaly region. J. Geophys. Res., 116, A02307.

Scotto, C., Pezzopane, M., 2011. Automatic scaling of polar ionograms. Antarctic Science. 24(1), 88–94. doi:10.1017/S0954102011000587.

Shi, J. K., Wang, G. J., Reinisch, B. W., Shang, S. P., Wang, X., Zherebotsov, G., Potekhin, A., 2011. Relationship between strong range spread F and ionospheric scintillations observed in Hainan from 2003 to 2007. J. Geophys. Res., 116, A08306, doi:10.1029/2011JA016806.

Abdu, M,. A., Kherani, E. A., 2011. Coupling Processes in the Equatorial Spread-F/Plasma Bubble Irregularity Development, book chapter in "Aeronomy of the Earth's Atmosphere and Ionosphere", Springer.

Abdu, M., A., 2012. Equatorial spread F development and quiet time variability under solar minimum conditions. Ind. J. Rad. Sp. Phys., 41, 168-183.

Alfonsi, L., Spogli, L., Pezzopane, M., Romano, V., Zuccheretti, E., De Franceschi, G., Cabrera, M., A., Ezquer, R., G., 2013. Comparative analysis of spread-F signature and GPS scintillation occurrences at Tucumán, Argentina. J. Geophys. Res. : Sp. Phys.. 118 (7), 4483-4502. doi: 10.1002/jgra.50378.

Cesaroni, C., Scotto, C., Ippolito, A., 2013. An automatic quality factor for Autoscala foF2 values. Adv. Space Res. 51 (12), 2316-2321. doi: 10.1016/j.asr.2013.02.009.

Haridas, M., K., M., Manju, G., Arunamani, T., 2018. Solar activity variations of equatorial spread F occurrence and sustenance during different seasons over Indian longitudes: Empirical model and causative mechanisms. Adv. Space Res, 61, 2585-2592, DOI: 10.1016/j.asr.2018.02.040.

Wan, X., Xiong, C., Rodriguez-Zuluaga, J., Kervalishvili, G., N., Stolle, C., Wang, H. 2018. Climatology of the Occurrence Rate and Amplitudes of Local Time Distinguished Equatorial Plasma Depletions Observed by Swarm Satellite. J. Geophys. Res.: Sp. Phys., 123(4), 3014-3026, DOI: 10.1002/2017JA025072.

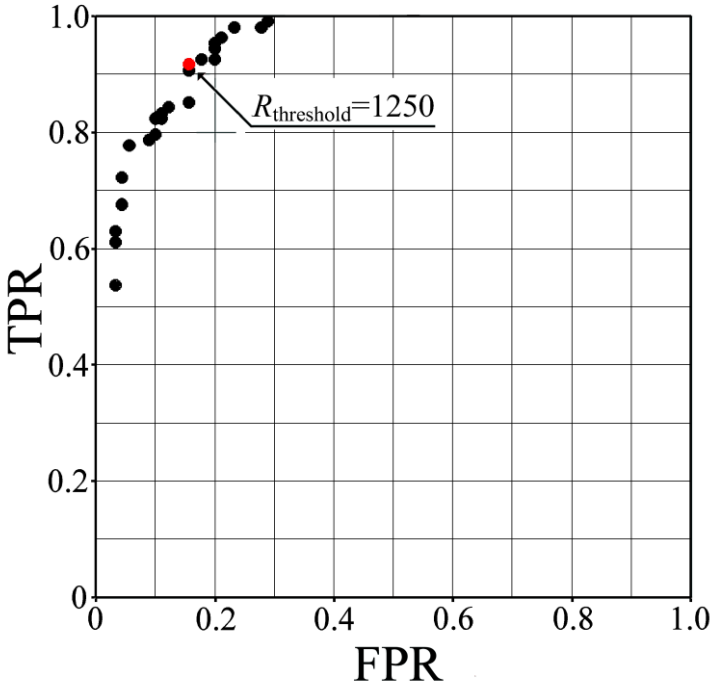

Fig. 1. ROC curve obtained from the data analysis. Based on this curve the $R_{threshold}$ was definitively set to 1250.

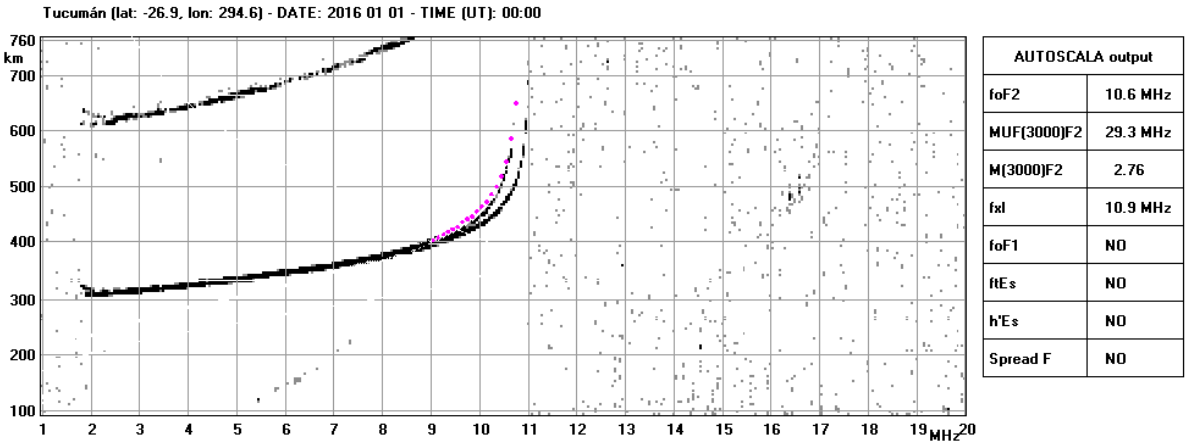

Fig. 2. An ionogram showing no spread-F features and correctly tagged (TN case). The *T* curve capable of fitting the typical shape of the F2 trace is included. In this case *R*=3602 and the parameters defining *T* are: *H=70*, *a=111*, and *A=13.0*.

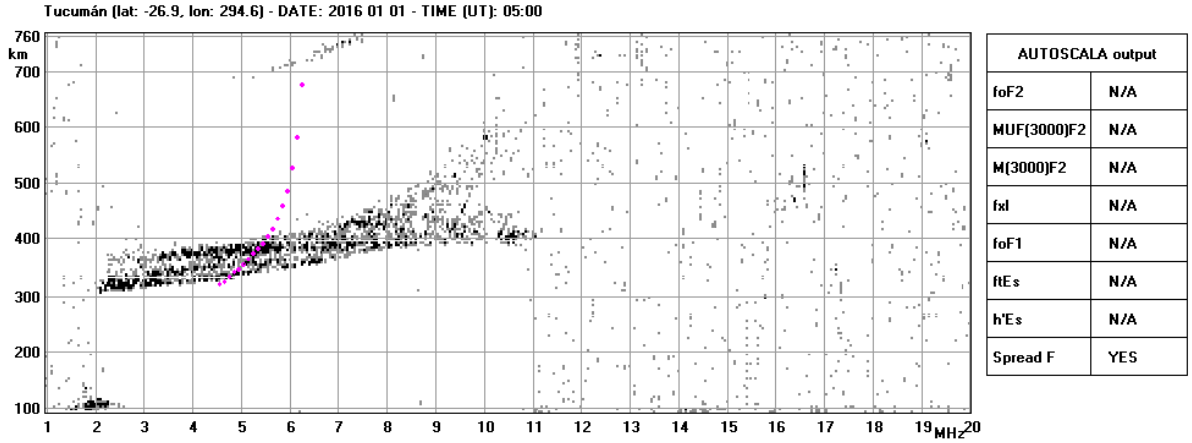

Fig. 3. An ionogram in which a clear case of equatorial spread-F is correctly identified (TP case). The *T* curve capable of fitting the typical shape of the F2 trace is included. In this case *R*=248 and the parameters defining *T* are: *H=51*, *a=66*, and *A=19.0*.

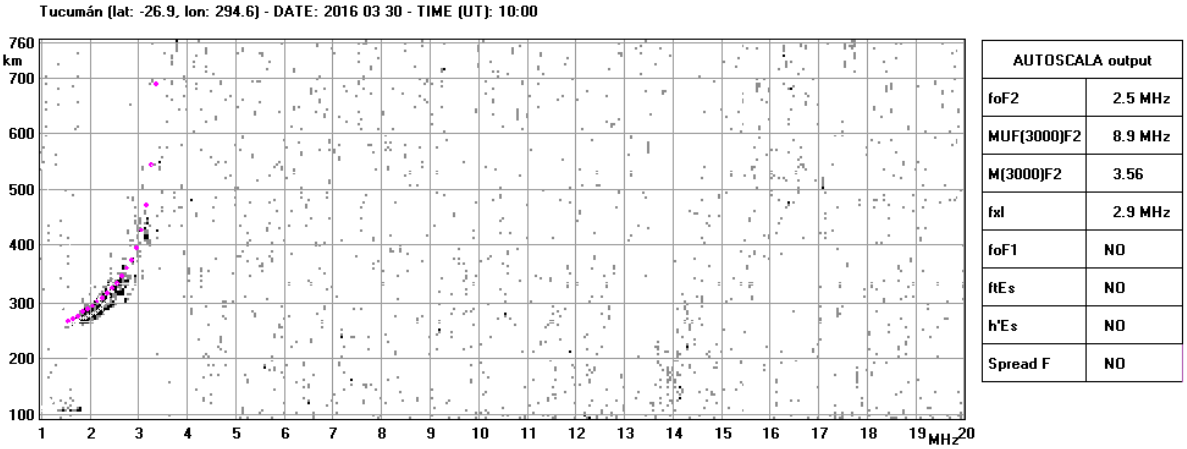

Fig. 4. An ionogram in which a spread-F trace is not detected by the routine, misleading Autoscala. An underestimated *foF2* value is given as output (FN case). The *T* curve capable of fitting the typical shape of the F2 trace is included. In this case *R*=1591 and the parameters defining *T* are: *H=39*, *a=36*, and *A=15.0* .

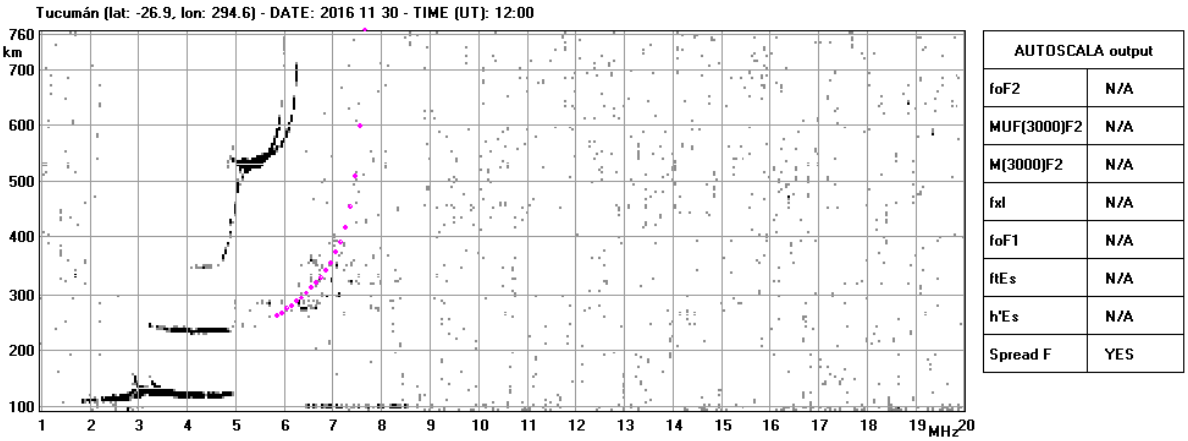

Fig. 5. An ionogram showing defined traces is wrongly tagged as featuring spread-F. The ionogram was discarded and no *foF2* value was returned as output by Autoscala (FP case). The *T* curve

capable of fitting the typical shape of the F2 trace is included. In this case R=784 and the parameters defining T are: $H=38$, $a=79$, and $A=18.0$.

Table 1. Results of TP, FN, TN, and FP obtained for the set I, after applying the routine with the decisional threshold set to 1250. The routine successfully detected a high percentage [TP/(TP+FN) = 81.55%] of spread-F cases. A small percentage of ionograms was wrongly tagged as being affected by spread-F [FP/(TN+FP) = 0.55%].

Ionograms with spread-F (607)		Ionograms without spread-F (7042)	
True Positive	False Negative	True Negative	False Positive
495	112	7003	39

Table 2. The performance of Autoscala in rejecting ionograms lacking sufficient information to establish an *foF2* value.

	Total	Scaled by the software		Not scaled by the software	
Not scaled by the operator	774	335	43.3%	439	56.7%
Scaled by the operator	6875	6740	98.0%	135	2.0%

Table 3. The performance of Autoscala in rejecting ionograms for which an *foF2* value cannot be established, either specifically due to spread-F or due to a general lack of information.

	Total	Scaled by the software		Not scaled by the software	
Not scaled by the operator	774	109	14.1%	665	85.9 %
Scaled by the operator	6875	6719	97.9%	156	2.3%