Rep. UAG 1051, Computer aided processing of ionograms and ionosande records,
World Dara Center A for Solar-Terrestrial Physics, National Geophysical Center, Boulder CO, USA
pp. 16-21, March 1998

MANUAL VERSUS AUTOMATIC COMPUTER PROCESSING FROM YEARS OF HOURLY DATA COMPARISON

Doc.Ing J.-C. JODOGNE,
Royal Meteorological Institute of Belgium, Av. Circulaire 3, Bruxelles, B-1180, Belgium, email: jcjod@oma.be

ABSTRACT

The ARTIST software outputs a near real-time set of parameters from the Digisonde 256 at the station Dourbes. These parameters are organized in monthly tables with the help of the ADEP software and are recorded in a file. ADEP is also used to manually validate or correct the scaling and output in other files. For each analysed parameter, and from the comparison of simultaneous determinations, a Kendall coefficient of concordance is computed and the difference (manual minus automatic) values are divided in classes to get the related statistical frequencies. Among the analysed parameters, only fmin and h'Es show less good concordance due partly to frequently missing E-region traces.

1. BUILDING THE DATA BASE

Monitoring of the bottomside ionosphere, by teledetection of the electronic density distribution, is one of the tasks of the Dourbes ionospheric station.

In routine use since 1984, a Digisonde 256 from Lowell University, USA, fed a computer running the ARTIST software (Reinisch & Huang,1983) and recording the raw and automatically processed data. The ARTIST outputs are edited and recorded, without any human intervention, producing the automatic station monthly parameters reports. Besides this, the ADEP software (Lowell) is used to manually validate, or correct, the trace and the calculated parameters and then produce the manual station monthly parameters reports (published and available). The data base encompasses the A files (automatic data) and M files (manual data).

2. PROCEDURE

For each analysed parameter, one takes the number, **m**, of manually determined values and the number, **a**, of automatically determined values to compare with the possible maximum number of values and get the respective relative numbers. The comparison is also made between these **m** and **a** numbers and the number, **c**, of simultaneous determinations by both methods. From these determinations, the difference between the "manual" value and the "automatic" one are computed. With this set of differences, one ranks in classes and computes the statistical frequencies for each class.

Finaly, a non parametric test, the Kendall coefficient of concordance (Siegel 1956), **w** ($0 < w < 1$), is produced from the simultaneous determinations. A high, or significant, value of **w** may be interpreted as meaning that automatic and manual scaling give essentially the same result in ranking the N determinations.

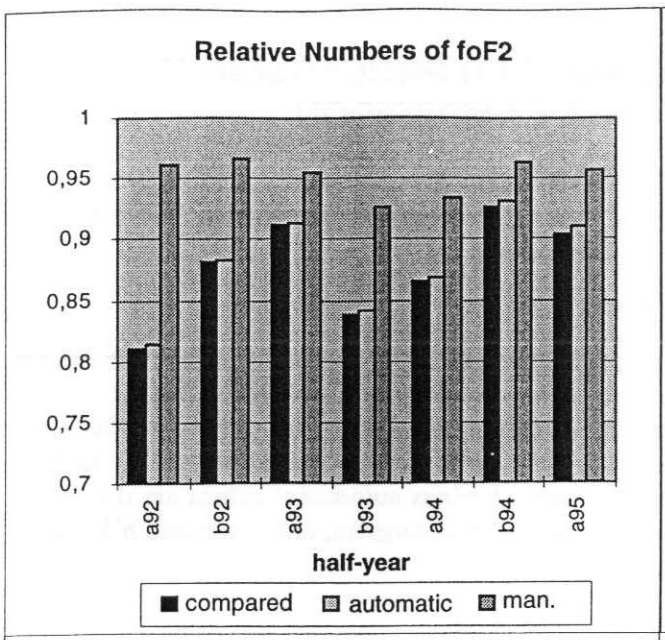
3. STATISTICS

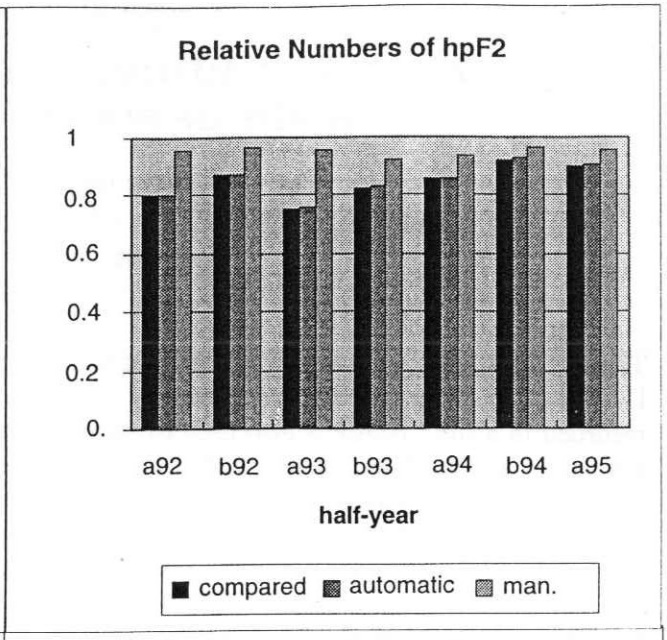
The data analysed were from January 1992 to June 1995 and were divided in half-yearly sets of (92a, 92b,...). The following ten parameters were selected:

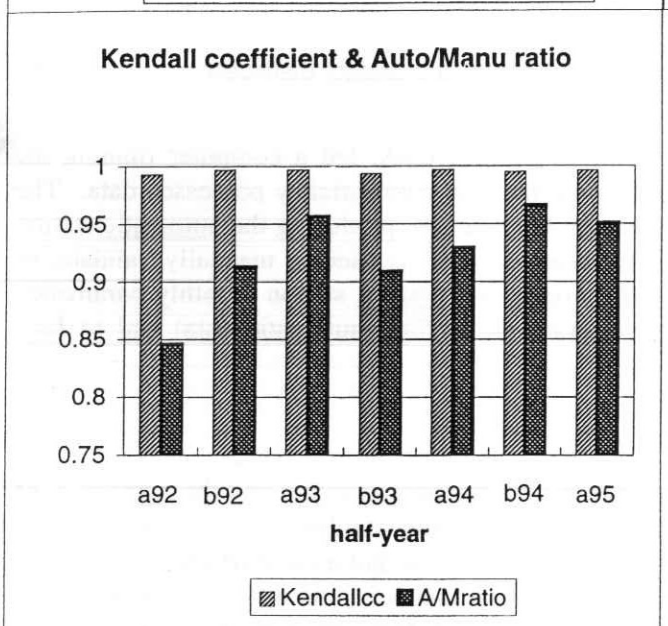
frequencies:	**foF2 , foF1, foEs, fmin, fminF, MufF2**
virtual heights:	**h'F, h'F2, h'Es**
M factor:	**M3000F2**

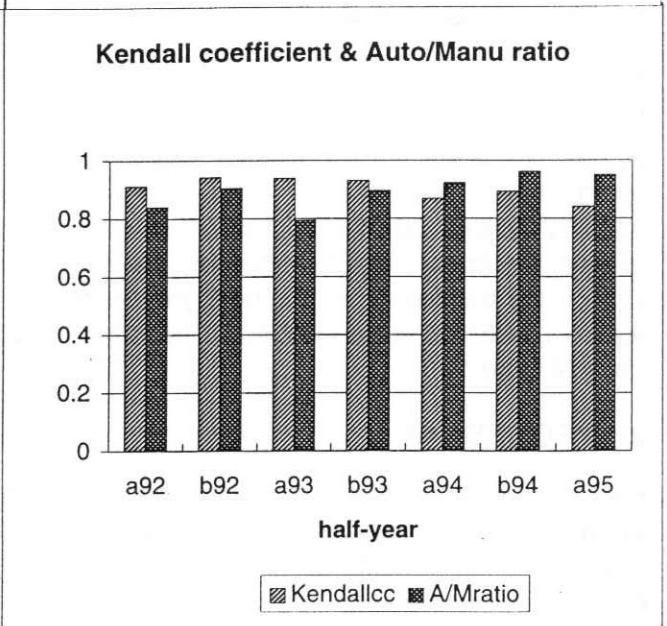
Other parameters were disregarded due to the existence of known biases in the comparison. For example, the automatic replacement of foE by the predicted value when there is no automatic determination or the fact that no manual determination of fxI was made before 1995.

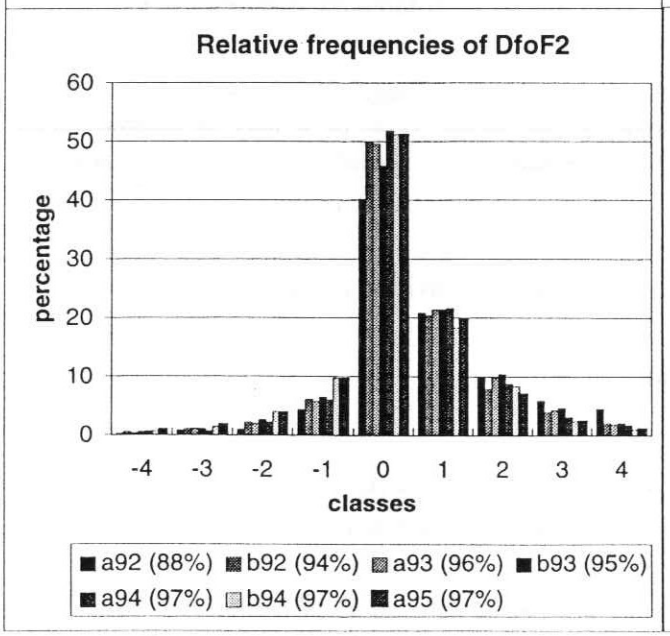
The maximum number for each parameter is 24 times the number of days, or 4344 for the first half of the year (or 4368 in the case of 1992) and 4416 for the second half of the year. The relative numbers always refer to these numbers. The total number of hours for the 3.5 years is 30648.











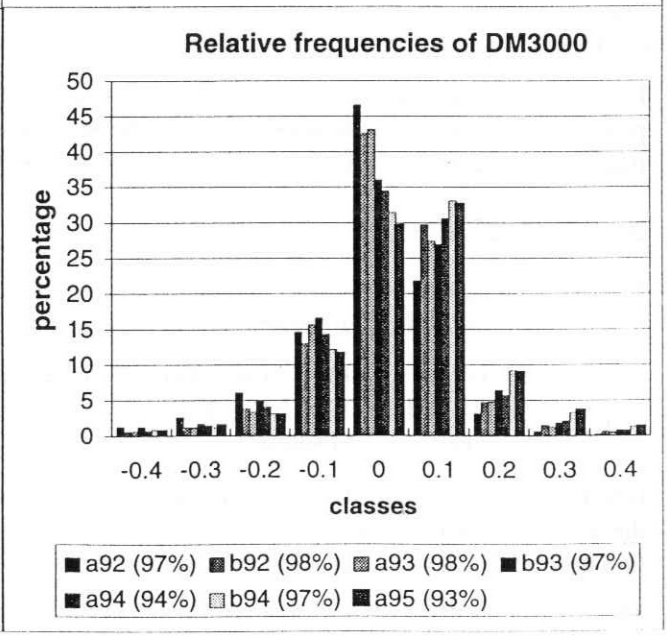
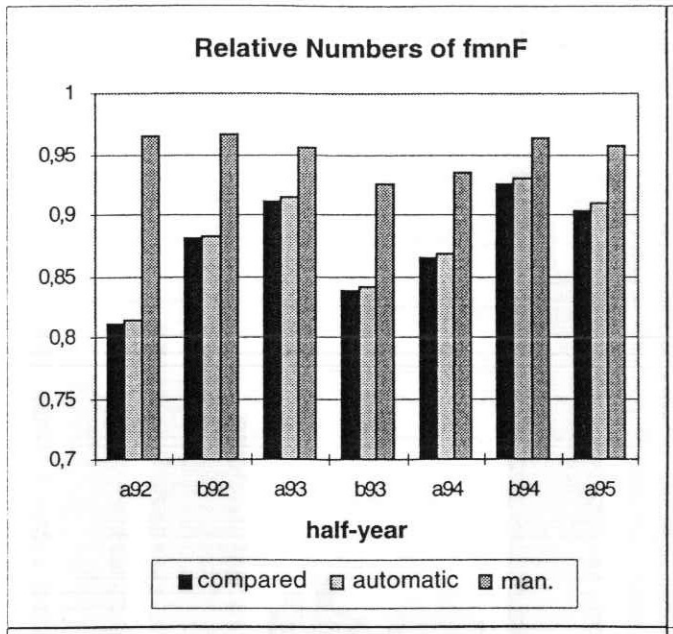
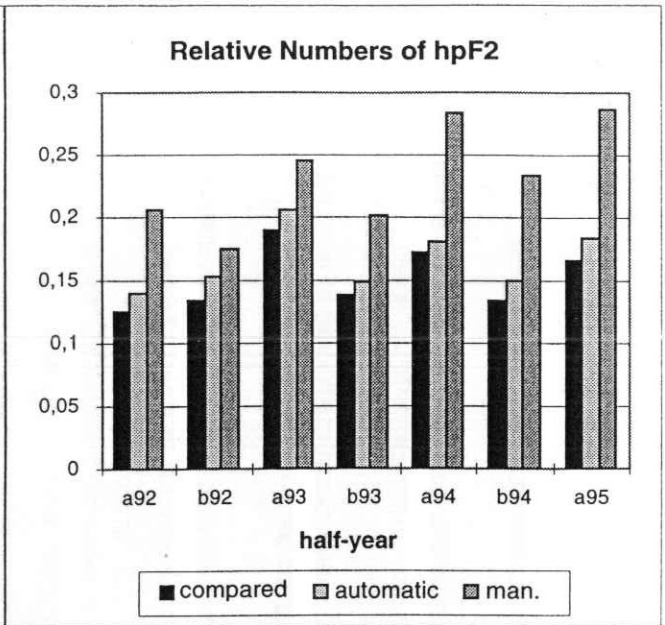
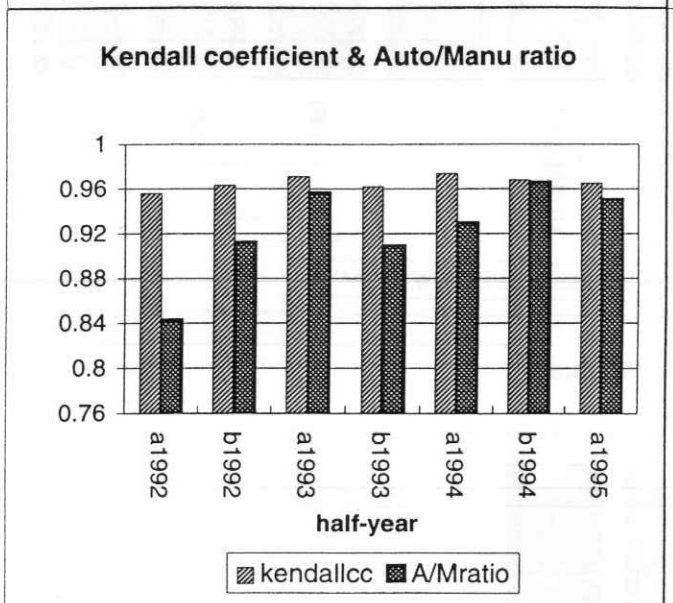


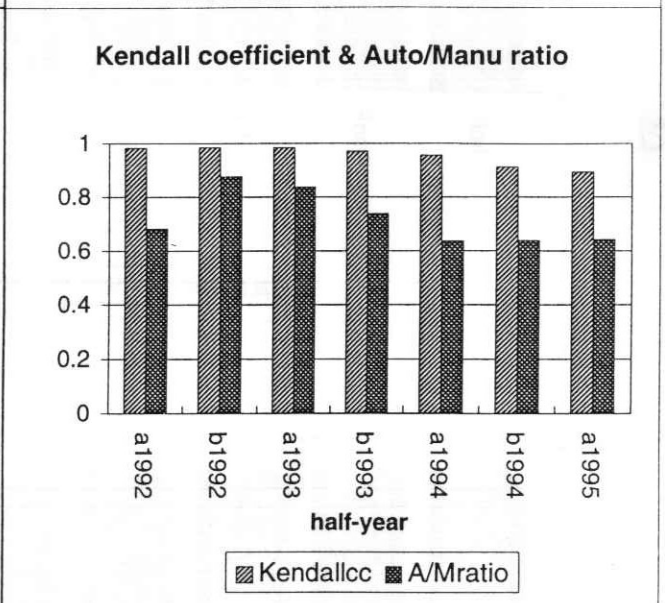
Fig.1a foF2 data

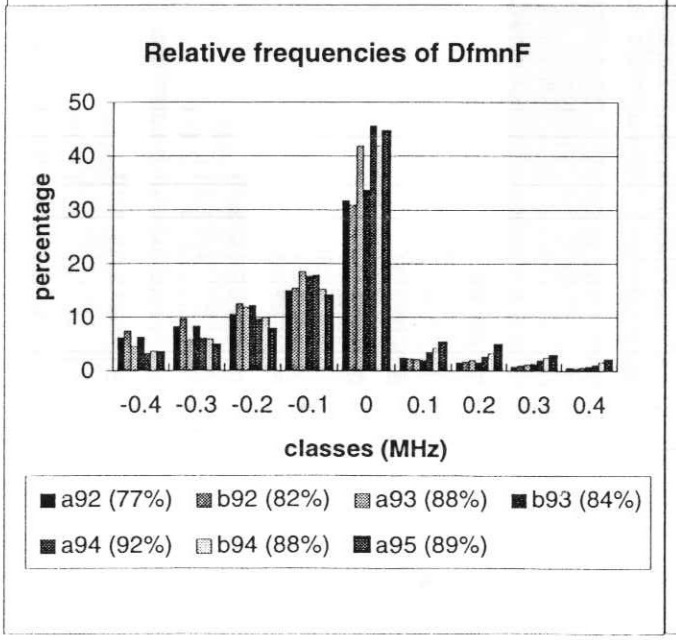
Fig.1b M3000 data











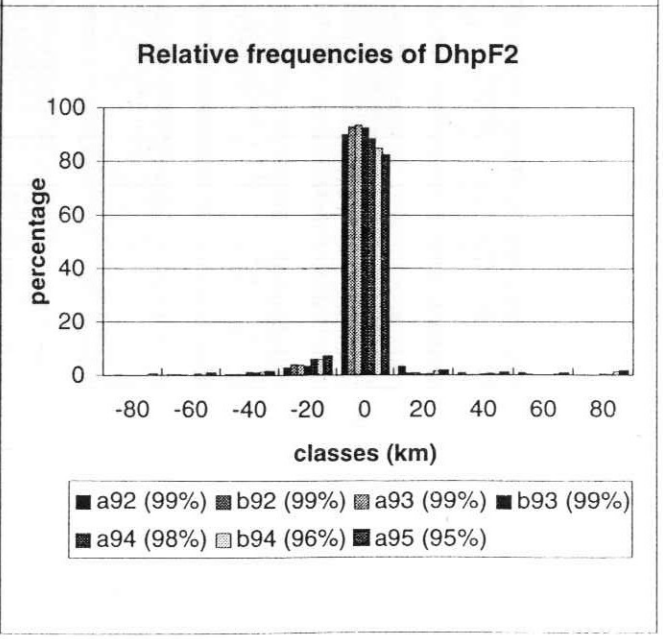


Fig.2a fmnF data

Fig.2b hpF2 data

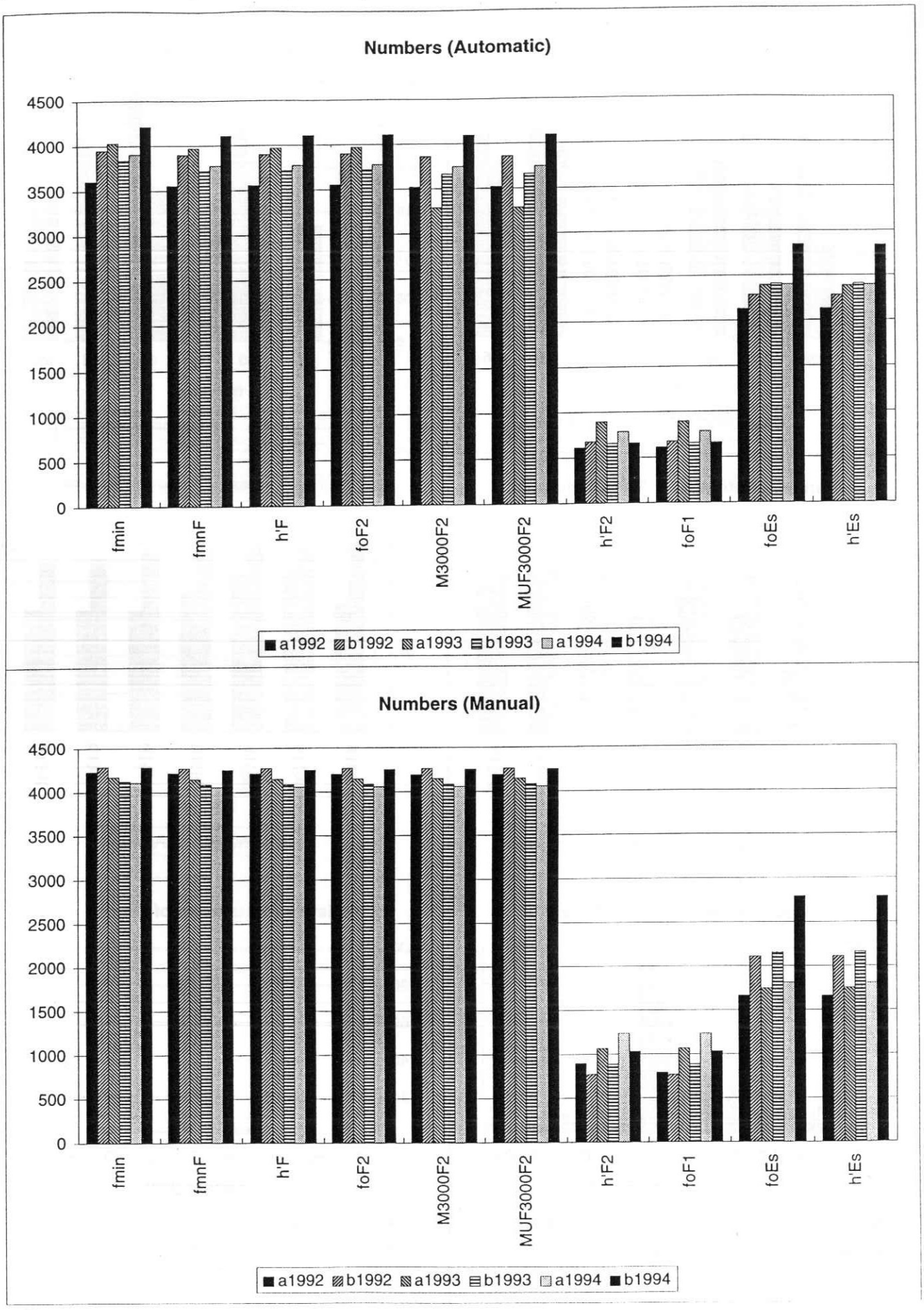


Fig.3 Numbers of automatic (top) and manual (bottom) determinations for 3 years

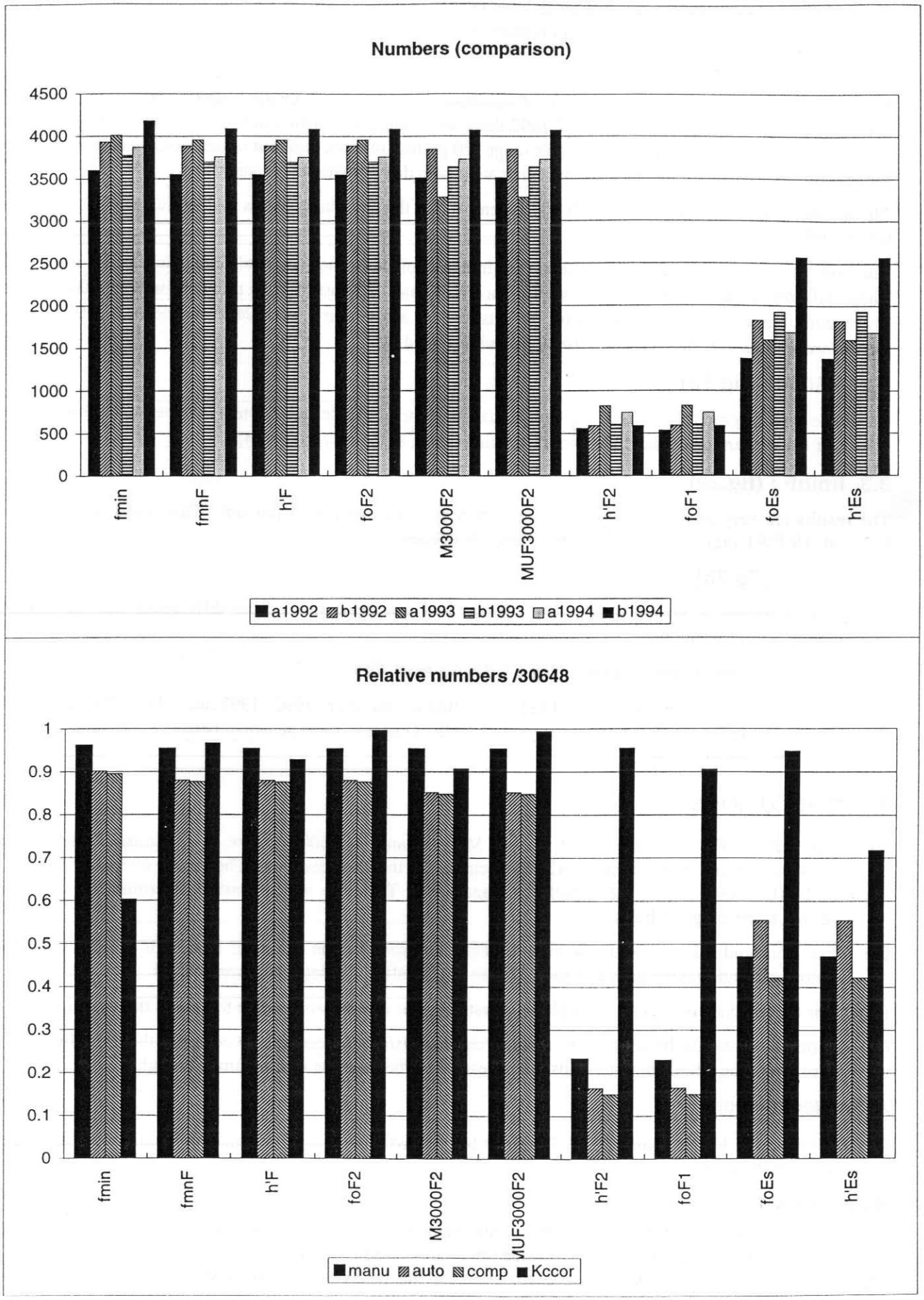


Fig.4 Numbers of comparisons (top) and relative numbers for all the data (bottom)

Power failure, ionospheric situation, breakdown or loss of data reduce the number of data especially for the 1st half-year of the year 1992. The main features are reviewed hereafter for some of the parameters.

3.1 foF2 : (fig.1a)

The top graph shows the relative numbers of simultaneous (rnberc), automatic (rnbera) and manual (rnberm) determinations. During the first half of 1992 there were some automatic data losses. Let us point out that the old version of ARTIST needs a 2 MHz range (20 points) to work and that noise could suppress the trace for several successive frequencies so that some automatic determinations are missing.

The middle graph displays the Kendall coefficient, **w**, for the automatic versus manual ratio and is excellent (over 0.99).

The bottom graph reveals that the differences (in MHz) of the values (manual minus automatic) are small. These differences are ranked in 9 classes for each half-year. Almost 50% of them are within ±0.05 MHz. The accumulative sum of percentages for the classes have a half-year label (a94(97%) means that 97% of all the differences for this half-year, for 1994, fell within ±0.45).

3.2 M3000 : (fig.1b)

The same type of graphs are displayed with a change of the left scale (from 0 to 1). Note the slightly higher values of the dispersion of the differences (in units) which still lies within ±0.45.

3.3 fminF : (fig.2a)

The results are very good, with the appearance of a more asymetrical dispersion. This could be due to the fact that ARTIST may sometimes use the double Es echoes.

3.4 hpF2 : (fig.2b)

This is an example of virtual height data. There is a low dispersion of the differences but the routine resolution is 5 km for one bin.

3.5 All parameters analysed : (fig 3a,3b,4a and 4b)

The first three figures (Figs. 3a, 3b, and 4a) are plotted for the years 1992, 1993 and 1994. The last figure (Fig. 4b) also includes the first half of the year of 1995. One sees some seasonal behavior although it is less prononced for manual data.

4. CONCLUSIONS

(i) For foF2, foF1, foEs, h'F, h'F2, fminF, M3000F and MUf3000F there is no statistical difference between manual and automatic ARTIST scaling for the Dourbes data. The Kcc is over 0.99 for foF2 and MUf3000F during the 7 half-years analysed. This also means profile determination in the F region starts on good basis.

(ii) For fmin and h'Es one must be circumspect. The contribution of fminE in fmin is not the same for automatic and manual determination. Noise may confuse the determination of h'Es.

(iii) For the analysis of special ionospheric situations, it is always worthwhile to look at the ionograms.

(iv) More time should be devoted to employ other statistical procedures to compare the data, but when looking at the details one may raise the question: is the machine or the human be right?

Acknowledgement

The author would like to thank Dr. F. Demeyer for discussions on statistics and for running the statistical software with the data.

References

[1] Reinisch, B. W. and Huang X. (1983), 'Automatic calculation of electron density profiles from digital ionograms. 3. Processing of bottomside ionograms', Radio Sci.,18,3,477.

[2] Siegel S., (1956), Nonparametric statistics for the behavioral sciences,McGraw-Hill, Int. Student Edition.