RESCUING IONOGRAM FILM ARCHIVES AT WORLD DATA CENTERS FOR THE IRI AND POSTERITY

J. W. Wright and R. O. Conkright

National Geophysical Data Center, NOAA, 325 Broadway, Boulder, CO 80303, USA

ABSTRACT

Some 200-million ionogram film images are held by the 4 World Data Centers, representing up to 60 years of global ionosphere and thermosphere weather and climatology. At best, about 25% of these have been reduced to a few "standard parameters", some of which are codified in the IRI. Fewer still have been reduced to electron-density profiles. Expertise in the rational analysis of the images is literally "dying off", while the film itself is at risk of decay. We advocate a special effort to rescue the information contained in these images, consisting of two distinct steps: (A) Scanning the film images into a suitable electronic format. (B) Extraction of numerical data by a pattern recognition process. We argue the necessity of action by the WDCs to avoid irretrievable loss of this precarious heritage. We offer ideas toward practical pattern-recognition software, and estimate the level of effort and costs required.

INTRODUCTION

An important part of the foundation of the IRI is usually identified with an empirical archive of hourly-values of ionospheric characteristics, some of which are easily available from the WDCs on two CD-ROMs (*Ionospheric Digital Database*, 'IDD', National Geophysical Data Center, Boulder). However, the foundation of *that* archive resides in some 200 million original film ionogram images, likewise held by the WDCs. More than 60 years of global ionospheric weather are documented in the film archives, with at least four times the temporal resolution of scaled hourly data. Many phenomena (e.g. Atmospheric Gravity Wave or 'AGW' effects) are aliased or escape notice altogether in hourly scaled data, but are well represented on the films. We seek to initiate some consideration of the problems of converting these resources into modern, machine-readable form, followed by reduction to physical quantities. Many benefits to the IRI, and generally to geophysics, will result from this effort. Examples are:

a) A four-times improvement in time resolution (from the present hourly data), and a factor of 12 on World Days since 1957;
b) Development of new parameters, including a Spread F climatology (Wright 2002); a "Sunrise Index" (d(foF2) /dCosχ; χ = solar zenith angle) at sunrise, serving as proxy for thermospheric [O]/[N2]), (Wright and Conkright, 2001); and extensive detail on Sporadic E;
c) Parameters from N(h) inversion, including hmaxF2 (with consequent thermospheric wind inferences), Total Electron Content, and an AGW database.

SCANNING IONOGRAM IMAGES

In the insert of Fig. 1 at lower-right, a typical ionogram image from 35mm film has been scanned by a simple desktop scanner (*Nikon* model LS3510, *ca.* 1994, used at ¼ of its maximum resolution of 2540 pixels/inch). A photographic reproduction of the original ionogram would be indistinguishable at this resolution. Analog ionosondes directly produced the originals, by photographing an oscilloscope range display on the slowly moving film during a frequency 'sweep'. Range markers (horizontal lines) and frequency markers appear, more or less dependably on each frame, followed by a station code, date and time. The photographic quality varies widely. The wanted data appear as irregular "traces" against the marker grid. Noise is a highly variable background from many sources. Human interpreters of these images used experience and theoretical understanding to reduce the images to numerical and graphic data. They worked at about half of the actual magnification provided by the main part of Fig. 1, obtained from the digital insert by a 20-fold zoom. At this magnification the individual pixels can just be seen. One pixel is about 1.67 km in echo range, and about 0.38% in radio frequency, comparable to the resolution attained by human scalers. Thus we may conclude that contemporary scanners maintain sufficient resolution for our purpose, in standard operation, to capture the frequency and height precision available in analog ionograms.

Typically, one or two 1000-foot rolls of 35mm microfilm comprise the output of one ionospheric observatory for one month's observations. Each such roll may contain about 3000 images, of special frame size (1 to 2 inches in length).

The images in Fig. 1 preserve a 'grayscale' quality of the original ionograms, rendered with 8-bit depth. Echo brightness might be thought to convey a measure of echo amplitude, of some potential value for purposes such as absorption, D-region modeling, solar flare effects, and irregularity diagnostics. However, the dynamic range of film is rather limited, and amplitude calibration problems were considered prohibive. For main-stream applications where the presence or absence of echoes was the dominating consideration, analog ionosonde designers often took one or more of the following specific steps to suppress recorded echo dependence on amplitude: (a) in the receiver, one or even two stages of fast signal differentiation (d(ampl.)/dt) and clipping utterly destroyed any surviving echo amplitude information. The steepened echo leading edge improved 'virtual height' precision also; (b) echoes were caused to blank, rather than to brighten the range-delay CRT display (as evident in Fig. 1); (c) film of deliberately low dynamic range was used.

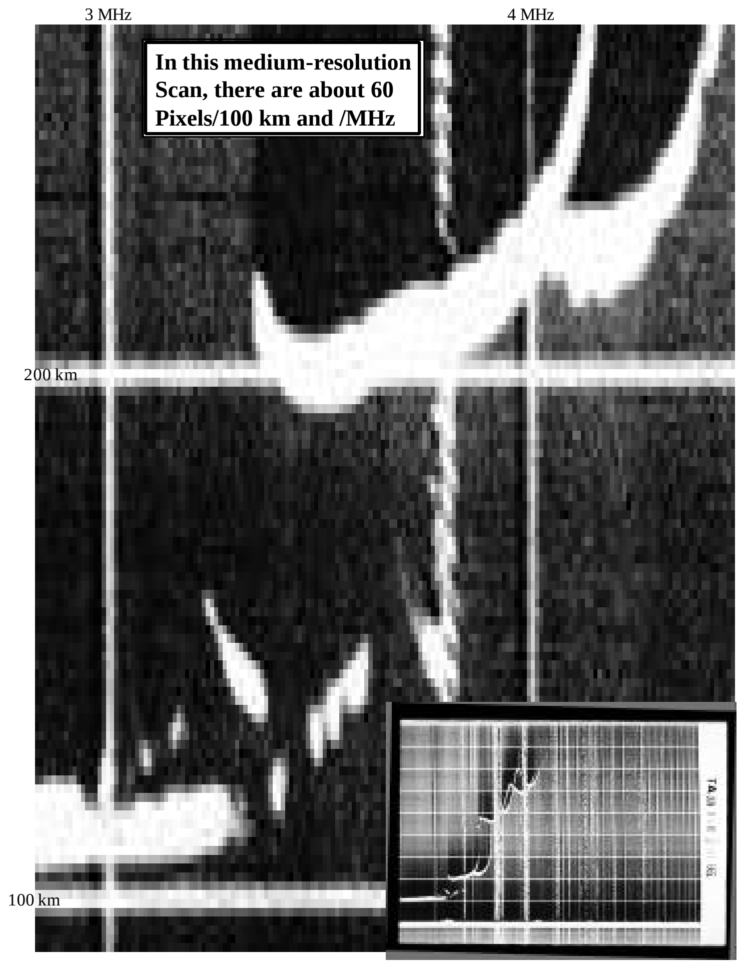

Fig. 1 (Insert) Grayscale scan of analog ionogram including Station, Date, Time stamp. (Main part) Zoomed part of E, lower F-regions showing resolution achieved.

Thus scanned grayscale images such as Fig. 1 preserve brightness information that is of little value. The factor of 8 in excess file size can be avoided, or can be more productively applied toward greater resolution, by representing the scan with (black-white) one-bit per pixel only.

Figure 2 is from an ionogram scanned in this manner, and with black and white interchanged. The scanning was kindly performed as a demonstration using modern industrial equipment by *Ronsin Imaging* (www.ronsin.com). They used a scanning resolution of 300 dpi, resulting in a '.tif' file size of 660 kB. Individual pixels may be just barely visible in the zoomed insert of Fig. 2. In a step not shown here, *Ronsin* applied a commercial software package (*TMS Sequoia's SCANFIX*) that discards stray unattached pixels; this reduced the file size to 390 kB, with no loss of useful information.

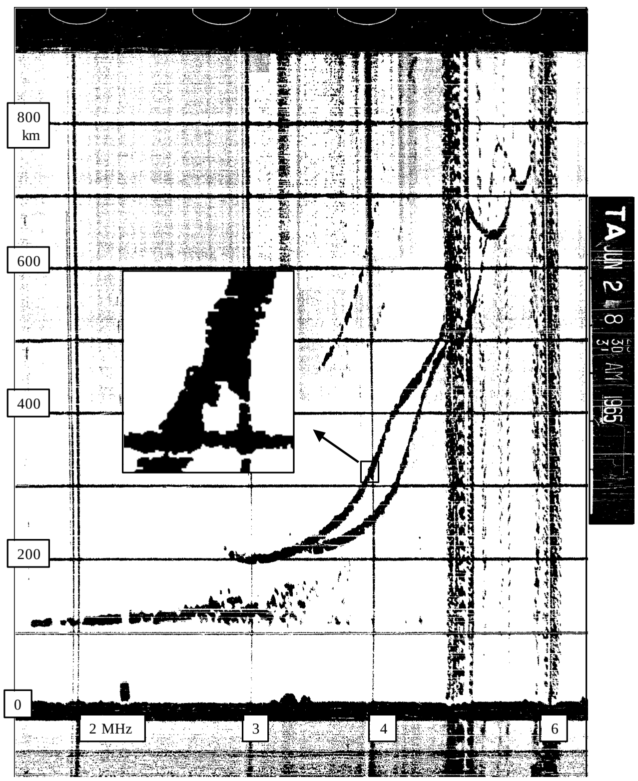

Fig. 2 Ionogram scanned 300 dpi with one-bit brightness depth (black or white pixels). File size 660 kB. Insert: portion zoomed 10x.

It is a step in the right direction, and technologically easy, if the original images were merely scanned and re-archived on DVD or other modern media. Final criteria for adequate resolution, noise rejection, etc., need to be worked out, but should not be excessively demanding. Considering the large size of the archive, it is essential that the scanning be highly automated: manual mounting of 1000 ft film reels is acceptable, but frame positioning must be fast and automatic.

Crude estimates of resulting archive size and cost are useful: At 400 kB per ionogram, the archive would occupy 80,000 GB or 107,000 CDs, or about 4300 4-ply DVDs. If service-bureau fees could be kept to $0.01 per ionogram, scanning would cost $2,000,000 plus media. Thus although scanning is practicable and would preserve the archive, it is not a trivial operation. Perhaps worse, the average scientist would be no closer to new data resources than at present. However, a further step, requiring significant development, can have profound effect on accessibility of the data in meaningful geophysical units, while probably permitting a reduction of the archive to very moderate proportions.

AUTOSCALING OF ANALOG IONOGRAMS

A far stronger motivation for this program can be generated by development of pattern-recognition procedures by which numerical data in geophysical units can be extracted from the images. This step would supplement the archive with an immediate geophysical resource of major proportions. Given the large number of images, this process must be entirely automatic and reasonably fast. The general problem may be broken down into five subtasks. Considerable progress has already been made on these, although in a somewhat different context, as will be explained later. In sequence, the subtasks are:

1. Identification of authentic echoes and their discrimination from noise.
2. Association of authentic echoes into “traces”.
3. Calculation of some numerical properties of each trace.
4. Grouping of traces according to general principles of ionospheric structure
5. Extraction of useful geophysical parameters.

Adaptable precedents for attacking these tasks should exist in ‘autoscaling’ methods developed for modern digital ionosondes, notwithstanding that such instruments may not even attempt to generate ionogram images as primary products, and their recorded data may include much more echo information than has been retained in the basic range-frequency coordinates of analog ionograms. Three such developments are known to us: The ‘ARTIST’ analysis for the digisonde (Galkin, et al. 1996) and the ‘SMARTIST’ analysis for the KEL ionosonde (http://www.arlut.utexas.edu/~engb/kel.html) are proprietary developments of their respective commercial sources, for which algorithmic details are not available. We can only surmise that they may address the same five subtasks, and must leave it to these sources to consider applying their concepts to the analog ionogram problem. The ‘DSND’ analysis system for the dynasonde, however, is explicitly structured according to the five steps above, is freely available as source code, and has been described by Wright and Pitteway (1997). We mention briefly how the DSND subtasks may be adapted to the restricted information content of scanned analog ionograms.

1. An efficient ‘coincidence detection’ scheme operates in real-time in the dynasonde: echoes are required to repeat within a small range tolerance δR' among a small number of repeated pulses at the same nominal radio frequency. Noise fails this test with high probability. For analog ionograms, repetition in time may be exchanged for repetition at nearby frequencies, with the additional feature that small, consistently-signed differences δR' are permitted between adjacent frequencies. Within a small interval δf/f (say, 3%, or 0.1 MHz near the center of Fig. 1) this process will perform the combined tasks of (a) recognizing authentic echoes; (b) rejecting noise; (c) yielding an average R' at average f, and (d) providing there an estimate of the trace slope, δR'/δf.
2. Echo association into “traces” in the dynasonde demands similarity of echo properties (amplitude, polarization, echolocation, Doppler) that are unavailable in scanned analog images. However, we suggest that sequences of [R', f, δR'/δf] triplets can be connected successfully into traces. In view of the characteristic shape of ionogram traces, it will be useful to interchange the roles of R', f as independent and dependent variables respectively, then using them together with 1/(δR'/δf) to define smooth traces by cubic splines.
3. We now have the basic numerical properties for each trace from steps 1 and 2. Forming trace mean R' and f are straightforward and useful for some downstream purposes. Trace “critical frequencies” (e.g., foE, foF2, fxF2, although not yet identifiable by name) can be determined accurately using a property of |1/(δR'/δf)| due to Paul and Mackison, (19xx): This quantity, considered as a function of f, extrapolates linearly to zero approaching layer penetration, whether from lower or higher frequencies, thus systematically estimating the R'(f) asymptote.
4. Testing “magnetoionic conjugacies” according to rigid rules in frequency permits distinguishing and naming penetration frequencies. For example, ‘O’ and ‘X’ penetrations must obey $f_o^{\,2} = f_x^{\,2} - f_x \bullet f_H$, where f_H is the local electron gyrofrequency. Conjugacy of penetration frequencies of course implies the associated trace conjugacies, but a further test is available for whole or partial traces: Group-range ratios, i.e. R'_0/R'_x, R'_o/R'_z, R'_x/R'_z, for the *same layer* and at conjugate frequencies, are only slowly varying with frequency (more strictly, with f / f_H) and are nearly independent of layer thickness and shape; they are, however, magnetic dip-dependent. For a particular site, a simple table of the ratios can be determined either from ionogram data, or from ray-tracing in model profiles, and then used in a trace-pair comparison to identify and confirm trace conjugacies.
5. The foregoing steps, while not as conclusive as their counterparts for precision digital recordings, should be sufficient to autoscale scanned analog ionograms. Remaining tasks include selection of traces for N(h) profile inversion, identification of sporadic E, quantification of spread F, rejection of simple “multiple echoes”, and interpretation of additional (implicitly off-vertical) traces as AGW signatures, etc.

It should be noted that a successful autoscaling process might improve considerably the economics of the preceding archival process. If a very high (e.g. 98%) success rate is achieved, the digital ionogram images could be replaced by the smoothed trace data (step 3 above), resulting in a much-reduced file size and media count. It might even be argued that the ionograms could be discarded altogether.

CONCLUSIONS

We have outlined the opportunities and requirements for a data-rescue effort demanded by the legacy of six decades of ionospheric monitoring by analog recording methods. The effort, consisting of film scanning to digital images and their reduction to geophysical data by a pattern-recognition process, will be rewarded by generation of a unique database of great contemporary value. Its implications for long-term environmental trends alone, probably justifies the endeavor. We consider it certain that if these efforts are *not* pursued, a significant program of major global cooperation, which peaked with the International Geophysical Year (1957) but which in fact was in place a solar cycle earlier and continues today, will have been done in vain.

ACKNOWLEDGMENTS

We are grateful to Ms. Kim Fisher of *Iimage Retrieval* (www.iiri.com), and to *Ronsin Imaging* (www.ronsin.com), for demonstrations of their 35mm film-scanning technologies.

REFERENCES

Paul, A.K. and D.L. Mackison, Scaling of the F–layer critical frequency from digital ionograms applied to observations during the solar eclipse on 26 February 1979. *J. Atmos. Terr. Phys.*, **43**, 221–224, 1981.

Galkin, I.A., Reinisch, B.W., Osokov, G.A., Zazobina, E.G., and Neshyba, S.P., Feedback neural networks for ARTIST ionogram processing, *Radio Sci.*, **31**, 1119-1128, 1996.

Wright, J. W., Quantifying Spread F by digital ionosondes, *Adv. Space Res.* 2002. This issue.

Wright, J. W. and R.O. Conkright, 2001 Prospects for an ionospheric index of neutral thermospheric composition, with space-weather applications *J. Geophys. Res*. **106**, 21,063-21,076, 2001

Wright, J. W. and M. L. V. Pitteway, Data acquisition and analysis for research ionosondes. In *Computer Aided Processing of Ionograms and Ionosonde Records*, *Proc. Session G5 at the XXVth General Assembly of URSI, Lille, France, Aug.28 - Sept. 5, 1996*. Ed. by Phil Wilkinson. Report **UAG 105**, NGDC, NOAA, 325 Broadway, Boulder CO 80303 USA. 1998.