The New ARTIST 5 for all Digisondes

Ivan A. Galkin and Bodo W. Reinisch

University of Massachusetts Lowell Center for Atmospheric Research
600 Suffolk Street, Lowell, MA 01854
Ivan_Galkin@uml.edu; Bodo_Reinisch@uml.edu

Abstract. Beginning in December 2007, UMLCAR is conducting a network-wide upgrade of the Automatic Real-Time Ionogram Scaler with True height (ARTIST), the ionogram autoscaling software developed for Digisondes. The new ARTIST Version 5 revamps classic, well-seasoned ARTIST 4 techniques for the modern computers and programming environments. It also implements new approaches developed to strengthen the ARTIST persona. Special attention has been directed to the characterization of the Autoscaling Confidence Level (ACL) and provision of the uncertainty bounds for the automatically scaled characteristics and derived electron density profiles in order to properly ingest them in assimilative ionospheric models. The new SAO.XML format for ionogram-derived data is used to report ARTIST 5 data. The paper describes progress in the ionogram autoscaling research at UMLCAR that has culminated into release of the ARTIST 5 upgrade.

INTRODUCTION

The Automatic Real-Time Ionogram Scaler with True height (ARTIST) [1] is an intelligent system developed at UMLCAR for automatic extraction of ionospheric specification data from digisonde ionograms [2,3]. Introduction of the ARTIST in the early 1980s was the single most influential advance in ionospheric sounding technology that had brought the ionosonde data outside of a narrow circle of experts into the realm of operational 24/7 space weather systems. Since then, the ARTIST went through several periods of regrettable lack of interest that major consumers of the ionogram-derived data endured in funding its further development. Recent renewal of attention to the ARTIST performance [4] is in part due to the advent of the Internet and establishment of the Digital Ionogram Data Base (DIDBase) at UMLCAR [5,6] that provide researchers and systems operators with instant public access to ionograms and profiles from the digisonde network. Realization of the evident risks associated with uninformed use of the autoscaled data in the assimilative models of the ionosphere stimulated studies of the ionosonde-related errors and uncertainties. In-depth analyses were conducted to specify common error types and sources [7]. Renewed attention was drawn to the automated quality assurance of the autoscaled characteristics [8] based on a variety of data sanity checks. The need for enhancing and improving the suite of ionogram-derived data products with respect to their accuracy has been clearly recognized. The ionospheric specification community is cautiously anticipating changes in the information flow between ionosonde data providers and users.

Clearly, each type ionosonde will develop its own autoscaling algorithms, and we cannot discuss in this short note the different approaches of other ionogram scaling algorithms that are in use or under development. While the main purpose of this note is to inform the users of digisondes and of digisonde data about the latest development of ARTIST, some of the ideas discussed in this paper may be of use to other algorithm developers.

FROM ARTIST 4 TO ARTIST 5

The need for significant changes in the existing ARTIST software stimulated yet another effort to develop the next generation ARTIST based on successful new software projects at UMLCAR: the pre-attentive vision model for extracting ionogram traces [9], the flagship digisonde data visualization and analysis software, SAO Explorer with DIDBase [6], and an intelligent image prospector [10] for the plasmagram images acquired by the Radio Plasma Imager (RPI) on the IMAGE satellite [11]. Even though existing Java code libraries at UMLCAR were largely reusable for the new ARTIST version 5, its development in 2002-2006 was mainly based on personal enthusiasm rather than a systematic, funded effort. However, the utility of the new ARTIST 5 was eventually recognized, trial installations began in 2006, and by the end of 2007 the critical mass of various advances in the ARTIST 5 project warranted its first official release for onsite operations.

The original ARTIST was a large FORTRAN application that comprises a wealth of logic considerations and a variety of *ad hoc* heuristics designed to replicate intrinsic intelligence involved in the process of visual ionogram interpretation. Many original design concepts have been preserved in ARTIST 5, with the exception of the trace-to-profile-to-trace optimized fitting technique, the centerpiece of the ARTIST version 4.5 [4], whose FORTRAN implementation remains to be translated to Java. A detailed description of the new algorithms developed in ARTIST 5 in comparison with its version 4 can be found in [12]. Here we limit our discussion to a brief summary.

Reduction of Echoes to Edgels

Ionospheric sounders, just like many other radar systems, look for the leading edge of detected echoes to evaluate the travel time of the signal. Many autoscaling algorithms, including ARTIST, replace detected echoes with their leading edge points (edge elements, edgels) for all subsequent ionogram analysis so as to reduce the computational complexity and needed computer resources. However, such classic, radar-style reduction to edgels has a well-known side effect of perceptual loss as illustrated in Figure 1.

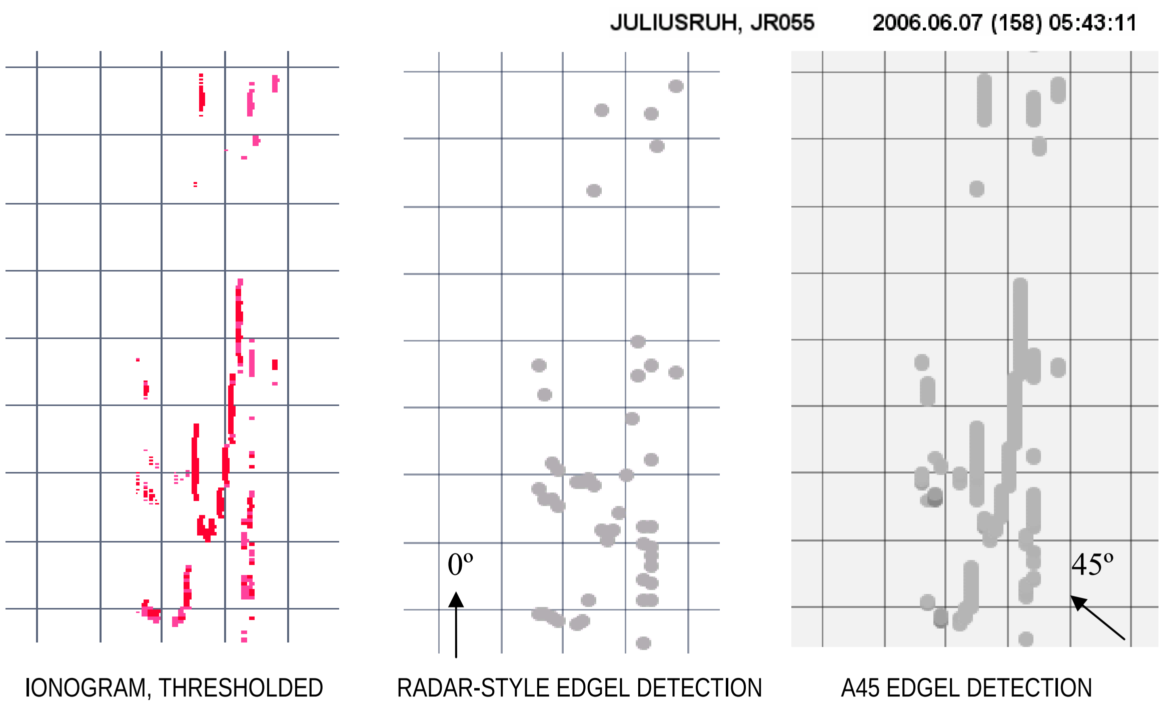

FIGURE 1. Edgel detection for ionogram traces with steep slopes. In the thresholded ionogram (left panel), each echo that occupies an interval of ranges has to be reduced to one point of its leading edge (edgel). Classic, radar-style detection (middle panel) slides the analysis window vertically within a single frequency. The A45 edgel detection algorithm slides the analysis window at 45 degree inclination, thus better representing the cusp shape.

The left panel of Figure 1 shows results of the echo detection in a digisonde ionogram recorded at Juliusruh, Germany, on June 7, 2006 at 05:43 UT. Summer time ionograms in the European sector at low solar activity are characterized by very short and steep F2 traces. The echo pulse usually fills several abutting range pixels on the ionogram, especially in the vicinity of the critical frequency cusps. When the ranges taken by such echo are collapsed to a single edgel (Fig. 1, central panel), the resulting dot pattern loses perceptual integrity of the cusps that was clearly visible in the original ionogram. Instead, ARTIST 5 employs a new A45 algorithm for echo-to-edgel reduction that scans the ionogram image not vertically, but rather at 45° angle as shown in the right panel of Figure 1.

Spread F Detector

Interpretation of ionograms recorded during spread F conditions requires a much higher degree of abstraction from the individual echoes comprising the spread trace. ARTIST 5 employs a detector of the spread F condition strength to adjust its processing algorithms accordingly. For moderate spread F levels, the abutting range pixels covered by the signal are represented by one edgel without finer analysis of the interval for overlapping echoes. For strong spread conditions, a smoothing median filter is applied to the ionogram and only one edgel of the strongest amplitude is allowed within the complete scan of each frequency. If the spread F severity is detected to be unacceptably high, the ionogram is labeled as inadequate for the analysis: most likely, it is impossible to find any clear vertical echo trace because the irregular electron density distribution does not allow defining a plasma frequency for a given height.

Pre-attentive Vision Model for Grouping Edgels into Traces

Grouping individual edgels into traces remains the greatest challenge of automatic ionogram interpretation. Human vision has an unsurpassed capability to extract fragmented, faint, intersecting, and dispersed traces from ionogram images. So far, the most successful computer solutions for this task rely on the availability of additional descriptive "tags" on the edgels, such as the wave polarization and the direction of arrival. Such tags simplify separation of echoes arriving via different wave propagation paths but appearing close to each other on the ionogram. The grouping of edgels can then be driven by the similarity of the tag values rather than more difficult visual saliency of the edgel sequences. While other types of distinctive echo "tags" have been actively researched for the purpose of such simplified but efficient edgel grouping [13], our hard-learned lesson taken from many years and locations of digisonde operations is not to expect reliable tagging for various real-life reasons. Even polarization tags can be wrong [14] as we further discuss below. Thus, we continued our work on the visual grouping algorithms that rely on the saliency of ionogram traces as described by the proximity and good continuation principles of the Gestalt perception [15]. Having this capability allows ARTIST to sustain reasonable quality of operations during the periods of instrument malfunction.

Pre-attentive vision models have been shown to work well for rapid detection of object contours in the field of view [16]. "ANNA", one of such models based on the Hopfield artificial neural network [9], was previously tested for detection of traces in digisonde ionograms and later applied in the CORPRAL software [10] for data prospecting in the voluminous plasmagram collection recorded by the RPI on NASA's IMAGE satellite [11]. We use the ANNA model in ARTIST 5 for extracting ionogram traces.

Use of X Polarization data

While previous versions of the ARTIST software used polarization tags available in digisonde ionograms mostly to rule out the X mode echoes, ARTIST 5 adopts X echoes to guide the ionogram interpretation process. Both the O and X-traces are extracted from the ionogram and evaluated for anchor cues. In particular, inspection of the critical frequency cusps in the F2 layer starts with fxF2 rather than foF2, because of its frequently more favorable representation in ionograms.

Fitting F2 layer cusps

The classic ARTIST concept of feature-driven fitting of hyperbolas to the ionogram image [1] is used to find the critical frequency cusps, foF2 and fxF2. There are several notable differences, though, that ARTIST 5 implements for the analysis of the F2 layer traces.

Separate hyperbolic fit

Using available O- and X-traces, hyperbolic fitting is done separately for foF2 and fxF2, and the derived hyperbolas are inspected for compatibility. Properly fitted cusps shall be separated by approximately one half the local electron gyrofrequency f_{ce}; if a mismatch is observed, the ionogram is re-evaluated to see if one of the cusps can be identified as poorly fitted and, if so, replaced with the replica of the other cusp, appropriately shifted in frequency. Figure 2 illustrates this approach by an example of the cusp fitting in the presence of restricted frequency (transmitter off) bands.

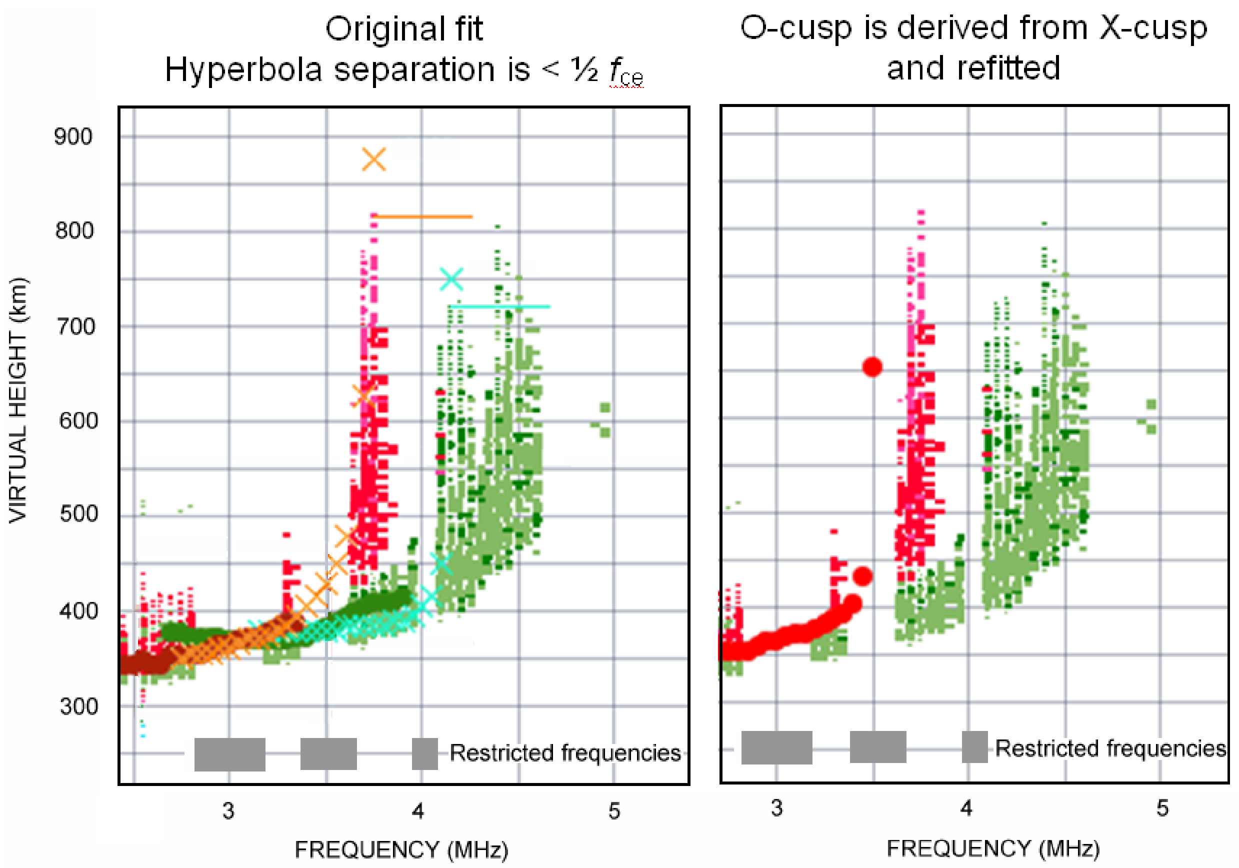

FIGURE 2. Hyperbolic fitting F2 layer critical frequency cusps in the presence of restricted frequency bands. (Left panel): Raw ionogram recorded at Millstone Hill is shown together with the extracted traces displayed by dark red (O) and dark green (X) dots. Original hyperbolas (crosses) are fitted to the trace tips. The resulting hyperbolas appear separated by less than one half the gyrofrequency f_{ce}, and also the X-cusp has a better fit than the O-cusp. (Right panel): The original O-cusp is discarded and its new shape is constructed by taking the X-cusp, shifting it in frequency by ½ f_{ce} and then shifting it in height to match the tip of the O trace. The final O-trace (shown as red dots) is the original O-trace whose tip is replaced with the newly derived O-cusp.

Both panels of Fig.2 show an example of a digisonde ionogram recorded at Millstone Hill observatory where transmission was restricted in the frequency intervals around 3, 3.5, and 4 MHz as indicated in Fig 2. In addition to the raw ionogram, the left panel of Fig. 2 shows traces (dark red dots for the O- and dark green dots for the X-polarization). The trace extraction algorithm was able to interpolate some gaps caused by transmission restrictions. Using extracted O- and X-traces, two model fits of the hyperbolic cusp to the ionogram were obtained, shown in the left panel of Fig.2 by orange (O) and cyan (X) crosses. However, the original fit of the O-hyperbola in the left panel of Fig. 2 was not accurate because of a missing section in the ionogram where ionosonde transmission was restricted. Fitted O- and X- hyperbola did not pass the check of their separation by ½ f_{ce}, thus triggering the fit quality test that determined X-cusp to be better. The original O-cusp was then discarded and replaced with a shape derived from the other polarization component, as shown in the right panel of Figure 2. The right panel of Fig. 2 shows the final O-trace (dots), which combines the original O-trace with the newly derived O-cusp.

Anomalies of F2 cusps

Anomalies of expected normal presence of both O and X cusps are detected early in the process of hyperbolic fitting. Normally, both cusps are present, but loss of either O- or X- cusp is commonly observed, in which case ARTIST 5 fits the available cusp and derives the other. Swap of polarization tags is typical for equatorial stations where signal propagation conditions (ionospheric tilts) may lead to reflection in the other hemisphere where the sense of the elliptical polarization is opposite. In this case ARTIST 5 swaps fitted hyperbolas for correct ionogram interpretation. Hardware malfunctions can cause mislabeling of X-data as O-data and visa versa, in which case ARTIST 5 reprocesses the ionogram without relying on the instrument tagging of the echo polarization.

Weighted hyperbolic fit

The hyperbolic fitting operation itself has been modified to favor hyperbolas with steeper gradient to avoid commonly observed cusp overshooting and ensure that during the spread F conditions the best fit follows the left boundary of the available echoes.

SPECIFICATION OF ARTIST UNCERTAINTY

Kalman-filter data assimilation models are expected to become the most promising technique for space weather applications following their success in meteorology and oceanography. New ionospheric assimilation models like the Global Assimilation of Ionospheric Measurements (GAIM) [17] differ from prior generation adaptive ionospheric models in that they analyze the uncertainty of the observational inputs before using them as constraints on the physical model drivers. ARTIST 5 software automatically determines the uncertainty of each provided *N*(*h*)-profile point [4]. In order to specify ΔN at each height, two boundary profiles, inner and outer, are determined (Fig. 3). The inner and outer boundaries reflect the uncertainties of the critical frequencies Δf_{cr} of each layer, the internal uncertainty of the starting height of the profile, and the uncertainties of the E valley model representation.

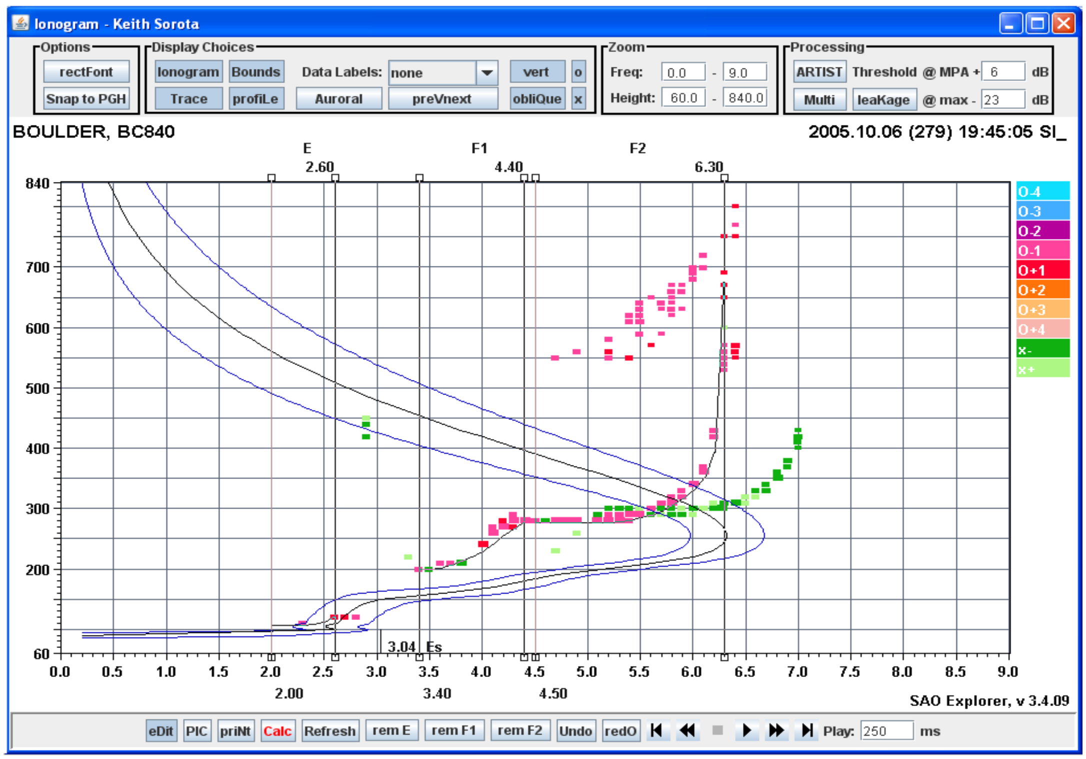

FIGURE 3. SAO Explorer ionogram display showing N(h)-profile with uncertainty boundaries.

The actual uncertainty values for the anchor points of the inner and outer boundaries are derived from the error histograms for each characteristic (*i.e.*, foF2, foF1) that are obtained by statistical analysis of differences between automatically and manually scaled values. The histograms are derived from a large number of scaled ionograms, individually for each station and separately for 3 possible levels of spread F (quiet, moderate, and severe). Once the error histograms are calculated, the lower and upper uncertainty bounds are selected to contain 95% of all histogram data, thus securing 95% probability that the true value falls within provided bounds.

ARTIST 5 AUTOSCALING CONFIDENCE LEVEL

As ARTIST 5 follows its logic steps in the interpretation decision tree, it can detect various anomalies that increase the uncertainty of the final result. For example, if individually fitted foF2 and fxF2 cusps are not separated by the expected ½ f_{ce}, thus triggering comparative analysis of the fit quality and subsequent replacement of one hyperbola with the other shifted by ½ f_{ce}, the overall confidence in the outcome of interpretation will be decreased. Other merit criteria affecting the Autoscaling Confidence Level (ACL) that ARTIST 5 reports in the data include the presence of O-polarization traces above foF2, large gaps in the F trace, multiple candidate traces in the F region, etc. One of the effective merit checks is inability of the profile inversion algorithm to derive a profile with the specified mismatch of the original traces and the traces restored from the calculated profile, which usually indicates a serious interpretation blunder.

We pay special attention to make sure that ARTIST 5 does not produce poor scaling results that are labeled with a high ACL rating. So far, our statistical comparison studies show improvement of autoscaling quality when low confidence records are excluded. Remarkably, only ~5% of ARTIST 5 records are excluded at locations where ionograms are known to be easy to interpret (mid-latitude, low interference observatories). We have seen up to 70% rejection of autoscaled results for ionograms taken at polar stations during severe spread F conditions.

Use of ACL is recommended for operational scenarios in which real time ionogram-derived data are driving models or prediction algorithms. While continuing development of ARTIST 5 aims at better accuracy of autoscaled data, it is beneficial to configure digisondes to higher measurement cadences and stricter confidence thresholds.

ARTIST 5 AUTOSCALING QUALITY

Figure 4 presents sample results of the ARTIST 5 uncertainty calculations made using the data collected by the digisonde station at Boulder, CO in 2005 (courtesy T. Bullett and R. Redmon). The uncertainty for foF2 was calculated from a total of 10,798 ionograms, while foF1 uncertainty was calculated from 2,029 ionograms analyzed. The results of the analysis shown in Fig. 4 demonstrate that for 95% of automatically processed ionograms the value of foF2 parameter determined with the ARTIST 5 was within the limits of (-0.3….+0.3 MHz) from the foF2 values determined by a human expert. Positive difference corresponds to the ARTIST 5 value being larger than the true value. Similarly, for 95% of the analyzed ionograms, the foF1 values suggested by ARTIST5.0 were found to be within the limits of (-0.4…+0.2 MHz) from the manually determined ones. The best frequency resolution for an ionogram measurement is defined by the unit step in the frequency sweep, which usually is set between 0.025 MHz and 0.1 MHz. A comparison of these numbers to the presented uncertainties of automatic ionogram scaling testifies for the high accuracy of ARTIST 5 scaling.

A similar comparison was performed for the data from other stations, Athens (courtesy A. Belehaki) and Jicamarca (courtesy O. Veliz), with the results summarized in Table 1.

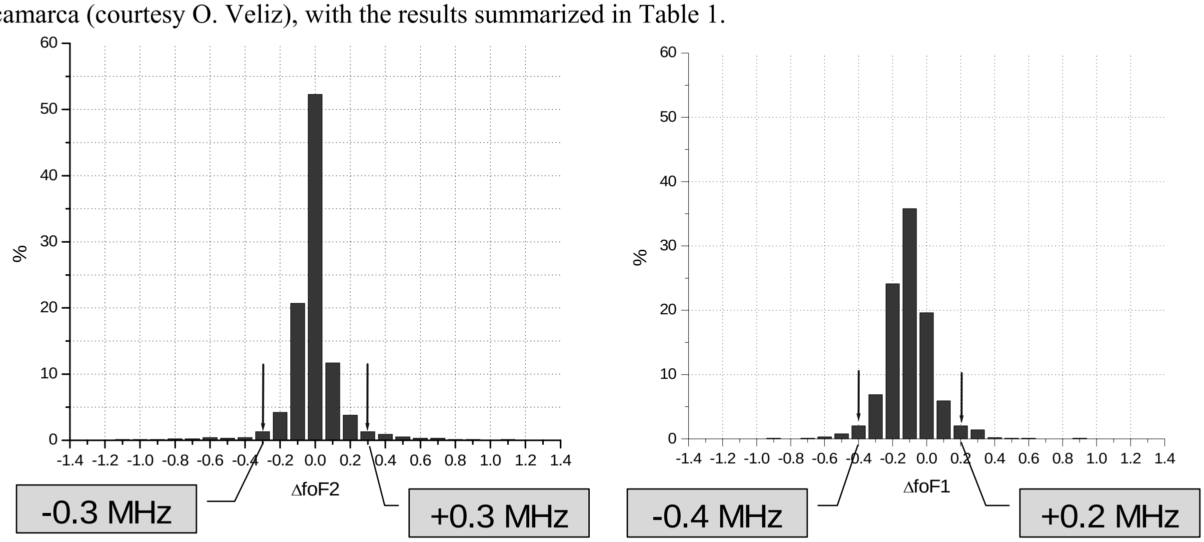

FIGURE 4. ARTIST 5 uncertainties at 95% for foF2 and foF1 parameters calculated for Boulder digisonde operation.

TABLE 1. ARTIST 5 uncertainties at 5% percentile

Station	ΔfoF2, MHz	ΔfoF1, MHz
Boulder	-0.3 .. +0.3 (10,798 cases)	-0.4 .. +0.2 (2,029 cases)
Athens	-0.3 .. +0.2 (6,645 cases)	-0.3 .. +0.4 (292 cases)
Jicamarca	-0.3 .. +0.4 (3,696 cases)	-0.6 .. +0.6 (184 cases)

SAO.XML

ARTIST software has used the Standard Archival Output (SAO) format [18] for storage of the ionogram-derived data since 1987. As the need for new data items to be reported in SAO files comes up, the current version of the SAO standard naturally undergoes appropriate modifications, each time spawning a network-wide update to the reading/writing software. The SAO format definitions underwent 5 notable revisions during 1990-2003; it is the painstakingly slow process of upgrading end user software that inspired the SAO design team to consider new approaches to the task of storing ionogram-derived data so as to allow forward-compatibility of the standard in its next revisions [19].

ARTIST 5 results are stored in the new SAO format version 5 called SAO.XML [19,20]. The SAO.XML format is based on the eXtensible Markup Language (XML), a popular choice for data exchange and long-term preservation. The XML combines extensibility with forward compatibility: addition of new data elements to future releases of the SAO.XML standard do not disrupt existing elements and do not require old software to be necessarily reworked in order to continue their operations within the original scope of their design. This is accomplished by distinctly marking each data element in the file, thus allowing unknown elements to be skipped. At the same time, SAO.XML continues to be human-readable, and even more user-friendly than previous SAO versions due to the addition of plain text descriptions to all stored items.

ARTIST 5 UPGRADE

The ARTIST software upgrade will be accomplished on per station basis as it involves its statistical and comparison studies versus existing versions 4.5 or 4. For each station individually, we will regression-test ARTIST 5 performance against manually scaled data to make sure the new version improves the autoscaling quality. These tests will provide the typical uncertainties of ARTIST 5 results to report in the output SAO.XML files. While the end users are upgrading their software to read/write new SAO.XML records, ARTIST 5 will continue to also generate output files in SAO 4.2 format.

SUMMARY

A new version 5 of the Automatic Real-Time Ionogram Scaler with True height (ARTIST) is available for upgrading the digisonde network. It revamps classic, well-seasoned ARTIST 4 techniques for modern development environments and implements new approaches developed to strengthen the ARTIST persona. While directly applicable only to the digisonde, we hope to draw attention of the key ionogram autoscaling experts to the ARTIST development trends. In particular, we have directed special attention to characterization of the Autoscaling Confidence Level (ACL) and provision of the uncertainty bounds for the automatically scaled characteristics and derived electron density profiles. Autoscaled data enhanced with these features are more appropriate for ingestion in assimilative ionospheric models. Also, the new SAO.XML format is used to report ARTIST 5 data. At the INAG Business meeting in New Delhi during the last URSI GA there was consensus that URSI recommend SAO.XML to as the new standard for ionogram-derived data exchange.

REFERENCES

1. B. W. Reinisch and X. Huang, Automatic calculation of electron density profiles from digital ionograms, 3, Processing of bottomside ionograms, *Radio Sci.,* **18**, 477-492 (1983).
2. K .Bibl and B. W. Reinisch, The universal digital ionosonde, *Radio Sci.,* **13**, 519-530 (1978).
3. B. W. Reinisch, Modern Ionosondes , *Modern Ionospheric Science* edited by H. Kohl, R. Rüster, and K. Schlegel, European Geophysical Society, 37191 Katlenburg-Lindau, Germany, 440-458, 1996.

4. B. W. Reinisch, , X. Huang, I. A. Galkin, V. Paznukhov, and A. Kozlov, Recent advances in real-time analysis of ionograms and ionospheric drift measurements with digisondes, *J. Atmos. and Solar-Terr. Phys.,* **67**, 1054-1062, (2005).
5. I. A. Galkin, D. F. Kitrosser, Z. Kecic, and B. W. Reinisch, Internet access to ionosondes, *J. Atmos. and Solar-Terr. Phys.,* **61**, 181-186, (1999).
6. G. M. Khmyrov, I. A. Galkin, A. V. Kozlov, B. W. Reinisch, J. McElroy and C. Dozois, Exploring Digisonde Ionogram Data with SAO-X and DIDBase, *Radio Sounding and Plasma Physics,* AIP Proceedings # 974, (2008).
7. T. W. Bullett, *Ionosonde Error Analysis*, January 7, 2004, NOAA/NGDC/WDC-A, ts.
8. L. F. McNamara, Quality Figures and Error Bars for Autoscaled Vertical Incidence Ionograms. Background and User Documentation for QualScan V2007.7, *Boston College Institute for Scientific Research Technical Report*, Chestnut Hill, MA, (2007).
9. I. A. Galkin, B. W. Reinisch, G. A. Ososkov, E. G. Zaznobina, and S. P. Neshyba, Feedback neural networks for ARTIST ionogram processing, *Radio Sci.,* **31**, 1119-1129, (1996).
10. I. A. Galkin, B. W. Reinisch, G. Grinstein, G. M. Khmyrov, A. Kozlov, X. Huang, and S. F. Fung, Automated Exploration of the Radio Plasma Imager Data, *J. Geophys. Res.,* **109(A12)**, A12210, doi:10.1029/2004JA010439, (2004).
11. B. W. Reinisch, D. M. Haines, K. Bibl, G. Cheney, I. A. Galkin, X. Huang, S. H. Myers, G. S. Sales, R. F. Benson, S. F. Fung, J. L. Green, S. Boardsen, W. W. L. Taylor, J.-L. Bougeret, R. Manning, N. Meyer-Vernet, M. Moncuquet, D. L. Carpenter, D. L. Gallagher, and P. H. Reiff, The Radio Plasma Imager investigation on the IMAGE spacecraft, *Space Sci. Rev.,* **91**, 319-359, (2000).
12. I. A. Galkin, G. M. Khmyrov, A. V. Kozlov, B. W. Reinisch, X. Huang and V. V. Paznukhov, The ARTIST 5, *Radio Sounding and Plasma Physics,* AIP Proceedings # 974, (2008).
13. J. W. Wright and M.L.V. Pitteway, Data Acquisition and Analysis for Research Ionosondes, *Computer Aided Processing of Ionograms and Ionosonde Records*, **UAG-105**, WDC-A for STP, 1-11, (1998).
14. M. Pezzopane and C. Scotto, Can the polarization tagging of the ionogram trace deceive the autoscaling methods? The Learmonth case, *Geophysical Research Abstracts*, **9**, 02671, (2007)
15. I. Rock and S. Palmer, The legacy of Gestalt psychology, *Sci. American*, **SCA9012**, 84-90, (1990).
16. H. Wersing, J.J. Steil, and H. Ritter, A Competitive Layer Model for Feature Binding and Sensory Segmentation, *Neural Computation*, **13**, 357-387, (2001).
17. R. W. Schunk, L. Scherliess, J. J. Sojka, and D. Thompson, Global Assimilation of Ionospheric Measurements (GAIM), *Radio Sci.*, **39**, RS1S02, doi:10.1029/2002RS002794, (2004).
18. R. R. Gamache, T. W. Bullett, Z. M. Zhang, B. W. Reinisch, and W. T. Kersey, ARTIST Data Editing and Printing Output Format - a Flexible Modular Database, in *Digital Ionogram Data Formats for World Data Center Archiving*, Proceedings of IIWG International Workshop, University of Lowell Center for Atmospheric Research, November 1989.
19. I. A. Galkin, G. M. Khmyrov, B. W. Reinisch, and J. McElroy, The SAOXML 5: New Format for Ionogram-Derived Data, *Radio Sounding and Plasma Physics,* AIP Proceedings # 974, (2008).
20. *SAO.XML 5.0 Specification*, http://ulcar.uml.edu/SAOXML/, University of Massachusetts Lowell, Center for Atmospheric Research, retrieved on January 8, 2008.

Acknowledgement
Part of this work has been supported by AF contract FA8718-06-C-0072.