Astronomical Data Analysis Software and Systems IV
ASP Conference Series, Vol. 77, 1995
R. A. Shaw, H. E. Payne, and J. J. E. Hayes, eds.

Data Reduction Software for SWAS

Z. Wang

Smithsonian Astrophysical Observatory, 60 Garden St., Cambridge, MA 02138

Abstract. The Submillimeter Wave Astronomy Satellite (*SWAS*) is an unique instrument that will observe from space several important spectral lines between 538 μm and 615 μm in wavelength. In anticipation of its operation starting 1995, a set of software tools has been developed, mainly as a layered package in the IRAF environment, to allow efficient inspection, processing, archiving and retrieval of the scientific data. In this contribution, we discuss the basic design concept and a few special features of these tools.

1. Introduction

As part of NASA's Small Explorers program, preparation is currently underway for the *SWAS* mission: a 0.6 m radio telescope in earth orbit capable of carrying out pointed observations in the submillimeter wavelengths. Its primary goal is to perform spectroscopic observations of the ground state or a low-lying transition in several important atomic and molecular species that are either difficult or impossible to detect from the ground. The *SWAS* Science Operation Center, established at the Harvard-Smithsonian Center for Astrophysics, is in charge of planning the scientific observations, preliminary calibration and analysis, and archiving all science data from the mission (see also Kleiner 1995)

In essence, *SWAS* is a self-contained small single-dish telescope working at very high frequencies (Tolls et al. 1994). It has a pair of heterodyne receivers that can be independently tuned, and a single acousto-optical spectrometer (AOS) with 1400 1 MHz channels of Thomson linear CCD digital output, enabling simultaneous observations of different lines. Despite similarities with ground-based single-dish millimeter and submillimeter telescopes however, *SWAS* is unique in several important aspects of its basic design and operation that require a software reduction package specifically tailored to its needs. In a departure from conventional software design for radio telescopes, we have implemented the *SWAS* software as a layered package in the IRAF/VOS environment, making extensive use of the host's powerful features such as the command interface, image I/O, vector processing, line fitting, and graphics capabilities. We have also incorporated the STSDAS Tables facility to help manage the relatively large amount of spectral header information. The main parts of this software have been coded in SPP and are now undergoing preliminary tests.

2. Science Requirements for SWAS Software

During each of its 97 minute orbits, *SWAS* will normally point at three to five different sources and perform repeated integrations at on- and off-source positions. There will be spectroscopic chop observations for compact sources, standard nod observations for extended objects, as well as mapping mode observations. The sampling rate of the spectrometer is one full spectrum every 2 s. The actual on-off source data calibrations are performed on every 2 s spectrum in each of its two spectral bands individually (each band occupies one half of the 1,400 channels of the AOS). Moreover, because the Doppler effect resulting from the motion of the spacecraft itself needs to be corrected for the upper and lower sidebands in opposite directions, the single AOS can actually yield four different spectra in each integration.

The observations are naturally grouped into segments depending on orbit and target. Each segment consists of up to about 35 minutes of on-off source integrations, plus calibration spectra. Sources in need of longer integration times will be observed in multiple segments. This requires that the shifting (for Doppler correction) and co-adding of spectra be easy and flexible in the reduction software, not only for spectra within each segment, but also between segments.

3. Outline of SWAS Software Design

The *SWAS* data reduction software is designed to perform the following basic functions: (1) check data integrity, find time gaps and possible noise spikes, and flag bad data according to various diagnostics and ancillary ancillary (mostly instrument housekeep) information; (2) apply (ON-OFF/OFF-ZERO) and passband calibration to each of the 2 s spectra; (3) evaluate and apply wavelength calibration spectra, and derive and apply Doppler corrections to each individual sideband before co-adding; (4) perform similar processing to the broad-band (continuum) channels to establish flux calibration points; (5) recognize planet observations and perform beam centering and planet calibration calculations; (6) sort and partition the data in ways that can be easily handled by a system-wide database management tool; (7) provide a user-friendly means to sample, evaluate, and display the data at every stage of the reduction process; and (8) produce FITS data files, each containing co-added spectra of one or more segments.

Five new IRAF tasks have been implemented in a local package (currently named "*submm*") to perform the calibration in consecutive steps. It is intended to work as a data pipeline, with many adjustable parameters in each of the steps to accommodate actual data reduction needs. In normal processing however, a list of data files can also be run through these tasks in a batch mode with entirely preset parameters.

Although the raw data files are unaltered in the reduction, output data files will contain stamps recording the time of the reduction and name of the reducer. This allows different reduction procedures to be performed independently on the same dataset, and leaves open the possibility of re-applying all of the "standard" calibration procedures with different parameter sets at a latter time.

4. Several Special Features

Because of the large number of short-integration (2 s) spectra that *SWAS* is to transmit to the ground, and the fact that each of these spectra has its own set of attributes, we have designed the *SWAS* data format in a way significantly different from the conventional approach.

Both raw and calibrated spectra are written in 2-D *images*, loosely resembling an IRAF "multispec" spectral image file. However, only the information that is common to all spectra are written into the image header. Most variable parameters for each individual spectrum are instead stored in two "associated" table files. In other words, each row in the table file represents a set of attributes for a row in the corresponding spectral line image. This approach eliminates the burden of an excessively long header for the image file, and more importantly, allows structured, efficient, yet flexible operations to be performed on nearly all of the spectral parameters. Tests show that this scheme is working for our intended purposes and the processing speed is satisfactory.

Because spectra of different receiver bands and sidebands share much of the same information, the actual data storage in file is a set of 2-D images, with the third dimension being "band". Only a small number of columns in the header table file are band-specific. This avoids redundancy and improves the efficiency of data processing and storage. The associated image and table files are clearly identified by their naming convention, and are grouped according to observing segments.

We archive the raw and calibrated *SWAS* data in the same fashion. Changes to a different file format can be performed with an auxiliary task, but is necessary only to export the data reduction product to other external data reduction systems for analysis or comparison. The output data will be made available in the more conventional "header plus spectral data" fashion and written in FITS, using the IRAF multispec format.

In the calibration, there is also a problem of automatically selecting appropriate OFF and ZERO to pair with given ON and CAL spectra in order to minimize noise and artifacts. These selections can be complicated because of the continuous nature of *SWAS* data taking, and the different modes of observations (nodding, chopping, and mapping). We have adopted a simple "stacking" approach, in which the most "recent" available OFFs and ZEROs are pushed onto the stack for subsequent processing of the ONs and CALs. The advantage of this approach is that it can be applied to all observing modes with little modification. Several user-changeable parameters are built-in the processing schemes such that we may select the best way to perform the reduction depending on the exact nature of the data.

5. Some Practical Considerations

SWAS data will be relayed to the SOC on a daily basis. The estimated volume of science data alone would be 120 to 150 MB per day. This is not a huge volume by today's standard, but it does require a substantially automated pipeline system that can process bulk of the data in an unsupervised batch mode. Also, the possible need for re-processing raw data through the pipeline requires that the processing time to be relatively short. Our current design calls for a processing time of a few hours for 24 hours worth of *SWAS* data (assuming a dedicated Sparc-20 workstation with 64 MB memory).

The RF COMB spectra taken in-orbit are fitted for precise line center positions (accurate to 0.02 pixel). These positions are then used to check for possible sudden frequency changes (mode hops). A dispersion relation for each band is then derived from the COMB lines and is represented by four coefficients. Actual wavelength calibration is performed by interpolating the dispersion relation time-wise.

We are currently using both SPLOT and SPECPLOT in IRAF to display the data. Although these are adequate for showing the spectra in pixel coordinates, some modifications need to be made to incorporate the World Coordinate Systems of our data to display physical coordinates. We are also hoping to adopt the new GUI packages SPECTOOL or ASpect currently under development (e.g., Hulbert et al. 1995) for our use.

It is straightforward to convert the IRAF image and table files into FITS files for archival purposes. The current plan is to make these FITS files into CD ROM's as permanent archive and distribution media. A problem still being explored is how to search and retrieve part of the data from the archive. As IRAF itself does not have a complete database system, we tentatively decide to save the key parameters which are likely to be searched for in a commercial database management system (Sybase).

6. Outlook

Although there are a number of existing software packages for radio telescopes that all do an excellent job in performing versatile data reduction procedures, none provides the exact tools for the specific needs of *SWAS*. One the other hand, being a relatively small project, we can not afford to develop a reduction system from scratch. The IRAF environment provides a powerful, open platform upon which smaller packages like our own can be built, and they can take advantage of many of the existing tools from NOAO as well as facilities being developed elsewhere. More importantly, since IRAF itself is continuously being improved and enhanced, we are looking forward to being part of its large active user community and remaining updated in this ever-changing astronomical software environment.

Acknowledgments. My thanks to Eric Mandel of CfA, Phil Hodge of ST ScI, and Frank Valdes of NOAO for their help in software related questions. I also wish to thank the *SWAS* science team and *SWAS* SOC staff for their support.

References

Hulbert, S. J., Eisenhamer, J. D., Levay, Z. G., & Shaw, R. A. 1995, this volume, p. **??**

Kleiner, S. 1995, page this volume, p. **??**

Tolls et al. 1994, IEEE Proc. Vol. 2268