

An Algorithm for Removing Fauna Corrupted Data Pairs During BASS Data Reductions

20 August 2004

Prepared by

S. M. MAZUK
Space Science Applications Laboratory
Laboratory Operations

Prepared for

SPACE AND MISSILE SYSTEMS CENTER
AIR FORCE SPACE COMMAND
2430 E. El Segundo Boulevard
Los Angeles Air Force Base, CA 90245

Engineering and Technology Group

APPROVED FOR PUBLIC RELEASE;
DISTRIBUTION UNLIMITED

20040917 152



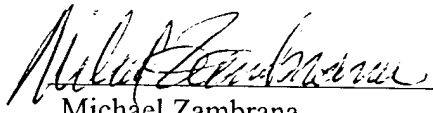
**THE AEROSPACE
CORPORATION**

El Segundo, California

This report was submitted by The Aerospace Corporation, El Segundo, CA 90245-4691, under Contract No. FA8802-04-C-0001 with the Space and Missile Systems Center, 2430 E. El Segundo Blvd., Los Angeles Air Force Base, CA 90245. It was reviewed and approved for The Aerospace Corporation by J. A. Hackweell, Principal Director, Space Science Applications Laboratory. Michael Zambrana was the project officer for the Mission-Oriented Investigation and Experimentation (MOIE) program.

This report has been reviewed by the Public Affairs Office (PAS) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nationals.

This technical report has been reviewed and is approved for publication. Publication of this report does not constitute Air Force approval of the report's findings or conclusions. It is published only for the exchange and stimulation of ideas.



Michael Zambrana
SMC/AXE

REPORT DOCUMENTATION PAGE			<i>Form Approved</i> <i>OMB No. 0704-0188</i>	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.				
1. REPORT DATE (DD-MM-YYYY) 20-08-2004		2. REPORT TYPE		3. DATES COVERED (From - To)
4. TITLE AND SUBTITLE An Algorithm for Removing Fauna Corrupted Data Pairs During BASS Data Reductions			5a. CONTRACT NUMBER FA8802-04-C-0001	
			5b. GRANT NUMBER	
			5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) S. M. Mazuk			5d. PROJECT NUMBER	
			5e. TASK NUMBER	
			5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) The Aerospace Corporation Laboratory Operations El Segundo, CA 90245-4691			8. PERFORMING ORGANIZATION REPORT NUMBER TR-2004(8570)-6	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Space and Missile Systems Center Air Force Space Command 2450 E. El Segundo Blvd. Los Angeles Air Force Base, CA 90245			10. SPONSOR/MONITOR'S ACRONYM(S) SMC	
			11. SPONSOR/MONITOR'S REPORT NUMBER(S) SMC-TR-04-17	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited.				
13. SUPPLEMENTARY NOTES				
14. ABSTRACT The purpose of this report is to document a procedure for selecting data pairs for analysis from Broadband Array Spectrograph System (BASS) data. BASS data are sometimes contaminated by aircraft or fauna entering the telescope beam during observations, obscuring the spectral signature of the object being measured. This procedure provides a mathematical method for identifying data pairs for exclusion from the subsequent analysis that may have been corrupted. BEST AVAILABLE COPY				
15. SUBJECT TERMS Infrared spectra, Data analysis				
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES 7
a. REPORT UNCLASSIFIED	b. ABSTRACT UNCLASSIFIED	c. THIS PAGE UNCLASSIFIED		
			19b. TELEPHONE NUMBER (include area code) (310)336-5614	

Contents

1. Introduction.....	1
2. Procedure.....	3
3. Calculations.....	5
4. Summary and Conclusions.....	7

Figures

1. Improvement in mean error for the AC-weighted signal for various combinations of pairs...	6
2. Comparison of filtered versus unfiltered AC-weighted data with error bars for the thermal infrared observations of Comet Neat (C/2001 Q4).....	6

Table

1. Pair Voting for the Three BASS Signals	5
---	---

1. Introduction

The Broadband Array Spectrograph System (BASS) is an infrared array prism spectrograph that covers the entire 2.9–13.5 μm spectral region simultaneously at a resolving power of 25–120, depending on wavelength.* The instrument is liquid-helium-cooled and uses two 58-element Blocked Impurity Band (BIB) linear arrays. End-to-end system performance (chopped) is NEP of $2 \times 10^{-14} \text{ W Hz}^{-1/2}$ at 10 μm and around $3 \times 10^{-14} \text{ W Hz}^{-1/2}$ at 4.8 μm . The circular entrance aperture is 2 mm in diameter. BASS can be configured for either $f/15$ or $f/30$ use depending on the application. The BASS instrument is often used at the Mount Lemmon Observing Facility (MLOF) in Arizona and the NASA Infrared Telescope Facility (IRTF) in Hawaii.

During observations at the MLOF and IRTF, it is not uncommon to encounter observing situations where something passes through the path between the telescope and the target object. While the angular extent of this path (referred to as the "beam") is small, it is not uncommon to have aircraft, birds, or insects cross through. Encounters with aircraft are generally few, but during temperate months the numbers of insects, birds, and bats can be significant enough to reduce the quality of a night's data.

The identity and spectral signature of the intruding objects are unknown, as is their total angular obscuration of the beam. Some encounters show such significant deviations that they can be seen in the channels monitored in real time by the observers, but smaller encounters may not be noticed. During post-processing the data containing the significant deviations can be excluded, but the smaller encounters would be missed. The data could be examined by hand for contaminated scans, but this is a tedious process at best and invites criticisms of "cherry-picking" the data. The procedure developed below represents a mathematical method of automating the exclusion of contaminated scans.

* Hackwell, J. A., D. W. Warren, M. Chatelain, Y. Dotan, P. Li, D. K. Lynch, D. Mabry, R. W. Russell, and R. Young, "A Low Resolution Array Spectrograph for the 2.9–13.5 μm Spectral Region," Proc. SPIE Conference 1235 on Instrumentation in Astronomy VII (1235), 171–180 (1990).

2. Procedure

BASS data are taken as \pm pairs in order to subtract out the sky background that is omnipresent in the infrared wavelength regime. One or both members of a pair can be contaminated by a foreign signal, so the pair has to be excluded from further analysis. A number of pairs are stored for a given data collection, and the data from these pairs are averaged together to obtain the final signal, referred to as a "sum file." The mean and standard deviations calculated for each wavelength from each group of pair data comprise the data saved in a sum file.

Now then we would like to calculate a new "sum" file from the data pairs, excluding the contaminated pairs. But the problem is that we don't know which one (or more) of these pairs may have been corrupted. The source of the corruption may be electrical or fauna in the beam, but the end result is a corrupt signal. The result of including a corrupt signal is to increase the size of the standard deviations in the resulting calculation.

The simple solution to this problem is to calculate a sum file for every combination of the pairs, and select the set of pairs that minimizes the errors. Thus, we have to evaluate $\sim 2^{N_{\text{PAIRS}}}$ number of sum files. Actually there are not quite this many since we have to have more than one pair to calculate statistics, so really we have $2^{N_{\text{PAIRS}}} - N_{\text{PAIRS}}$.

We are going to implement this using a vector mask to select which pair to use. So if we have three pairs, we set up a mask as [1,1,1] to use all three. To deselect a pair, say pair 2, we would use a mask like [1,0,1]. To generate these mask vectors, we can use bit operators on an integer number. For our three pair example we have:

0 = [0,0,0]	-> no pairs; don't use
1 = [0,0,1]	-> one pair; don't use
2 = [0,1,0]	-> one pair; don't use
3 = [0,1,1]	-> pairs 2,3
4 = [1,0,0]	-> one pair; don't use
5 = [1,0,1]	-> pairs 1,3
6 = [1,1,0]	-> pairs 1,2
7 = [1,1,1]	-> pairs 1,2,3

So we are only calculating for 3, 5, 6, 7, or 4 pairs total, not quite as onerous as it first looked!

We now have a large set of statistics calculated, so what are we going to do with it? Recall that the objective here is to find a set of pairs that reduces the error over the whole spectral range. Another way to say this is to remove pairs to find the set that has the smallest error.

The a priori information we are working from is that all the pairs are generated while viewing the same source and that the instrument is stable over the time of the measurement. We can make no

assumptions about the spectral signatures of either the source or the contamination. The only assumption is that the contamination is present in a limited subset of the available number of scans.

So using the standard deviation as a measure of the error, we need to figure out which pairs to exclude. Since the BASS data are spectral, we chose to average the standard deviation over wavelength to make the decision based on a scalar quantity. We could pick a single wavelength or a range of wavelengths, but with no prior information, the average is a good place to start. We do know that the real-time output is generated from a subset of the available wavelength channels, and we will leave the option of looking at a user-defined subset in the procedure. Further, we would like to maximize the number of pairs kept in the average, so as to increase the statistical fidelity of the answer. The process then is to average the standard deviation over the desired spectral interval for all possible pair combinations, reduce the combination space based on a user-input criterion of number of pairs to discard (e.g., discard 1 pair only), compare against the "all-pairs" case, and chose the combination with the lowest value.

The BASS data consists of three signals: the AC weighted, DC weighted, and DC signals. These are each handled independently in the calculation procedure, and the optimal selection of pairs is made for each signal. The final decision on pairs is made as a majority vote between the three signals, with equal weighting given to each signal.

3. Calculations

Data from observations of Comet NEAT (C/2001 Q4) taken 21 June 2004 from MLOF were used for testing. This particular observation was known to have a single bad pair based on the real-time signals monitored during the observations. The data consisted of 10 pairs, resulting in a total of 1013 possible combinations to be tested. The algorithm was allowed to test for discarding a maximum of 5 pairs.

The resulting 638 combinations for the AC-weighted error term are shown in Figure 1. here the y-axis shows the percentage difference in error as compared to using all 10 pairs in the calculation. The best pair combination here was pair index 546, which excluded pairs 5 and 6, as shown in the algorithm output of Table 1. The voting process between the three signals resulted in pair 6 being excluded.

Figure 2 shows the resulting filtered data in the thermal infrared region of the spectrum. The change in error bars is subtle in comparison, as would be expected for the small improvement obtained. Since the improvement decision is based on the wavelength-averaged error, an improvement in a noisy region of the spectrum will have significant impact on the final decision.

Table 1. Pair Voting for the Three BASS Signals. From these votes, pair 6 was selected to be discarded. (1 indicates a vote to keep; 0 indicates to discard.)

Pair Index	AC Weighted	DC Weighted	DC Signal
0	1	1	0
1	1	1	1
2	1	1	1
3	1	1	1
4	1	1	1
5	0	1	1
6	0	1	0
7	1	1	0
8	1	1	0
9	1	1	0

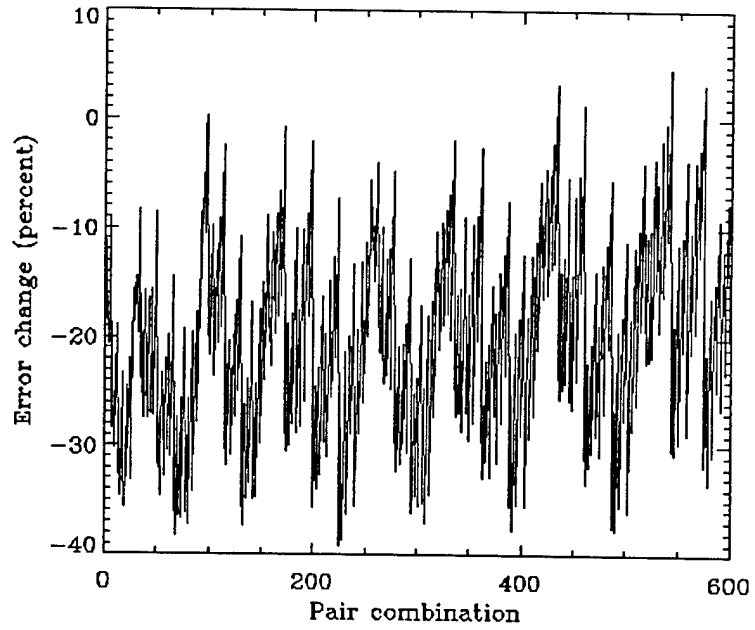


Figure 1. Improvement in mean error for the AC-weighted signal for various combinations of pairs. Here, the maximum improvement is obtained for pair index 546, which excludes pair 6 from the calculations.

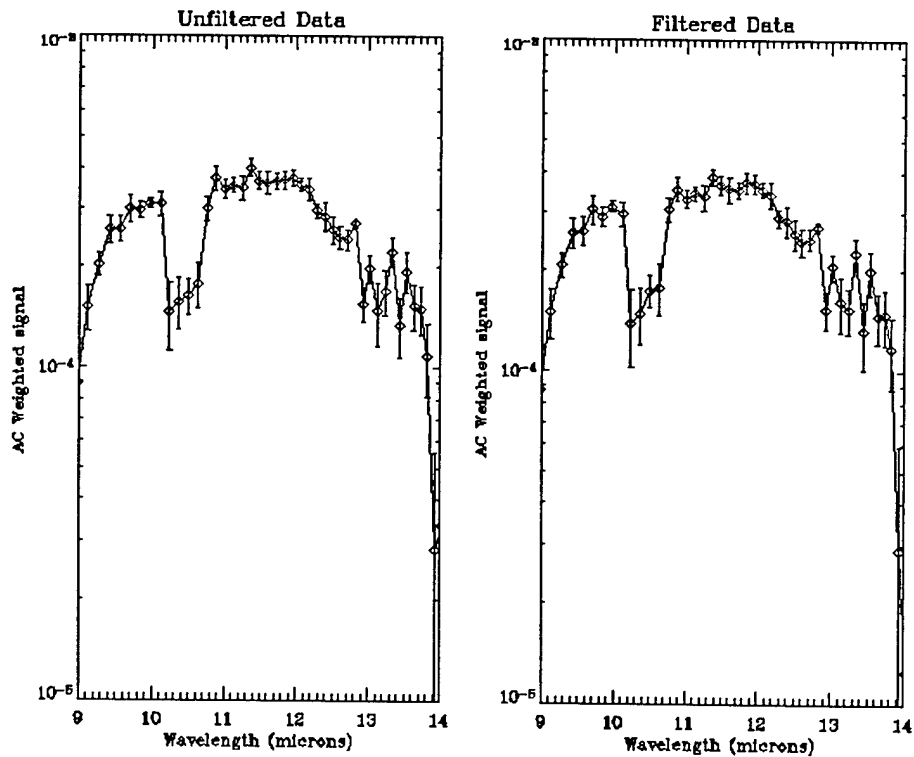


Figure 2. Comparison of filtered versus unfiltered AC-weighted data with error bars for the thermal infrared observations of Comet Neat (C/2001 Q4). These data were taken 21 June 2004 with BASS from MLOF.

4. Summary and Conclusions

A procedure has been developed to filter BASS pair data to mitigate the impact of pairs contaminated by aircraft or fauna. The procedure involves searching the entire combination space of all possible pairs. While this requires a significant number of calculations, the procedure runs with acceptable performance on a late-model desktop computer. This tool provides a new capability for the analysis of BASS data to improve the confidence in the results. Future work along these lines is anticipated in the development of alternative metrics for deciding which pairs to discard.

LABORATORY OPERATIONS

The Aerospace Corporation functions as an "architect-engineer" for national security programs, specializing in advanced military space systems. The Corporation's Laboratory Operations supports the effective and timely development and operation of national security systems through scientific research and the application of advanced technology. Vital to the success of the Corporation is the technical staff's wide-ranging expertise and its ability to stay abreast of new technological developments and program support issues associated with rapidly evolving space systems. Contributing capabilities are provided by these individual organizations:

Electronics and Photonics Laboratory: Microelectronics, VLSI reliability, failure analysis, solid-state device physics, compound semiconductors, radiation effects, infrared and CCD detector devices, data storage and display technologies; lasers and electro-optics, solid-state laser design, micro-optics, optical communications, and fiber-optic sensors; atomic frequency standards, applied laser spectroscopy, laser chemistry, atmospheric propagation and beam control, LIDAR/LADAR remote sensing; solar cell and array testing and evaluation, battery electrochemistry, battery testing and evaluation.

Space Materials Laboratory: Evaluation and characterizations of new materials and processing techniques: metals, alloys, ceramics, polymers, thin films, and composites; development of advanced deposition processes; nondestructive evaluation, component failure analysis and reliability; structural mechanics, fracture mechanics, and stress corrosion; analysis and evaluation of materials at cryogenic and elevated temperatures; launch vehicle fluid mechanics, heat transfer and flight dynamics; aerothermodynamics; chemical and electric propulsion; environmental chemistry; combustion processes; space environment effects on materials, hardening and vulnerability assessment; contamination, thermal and structural control; lubrication and surface phenomena. Microelectromechanical systems (MEMS) for space applications; laser micromachining; laser-surface physical and chemical interactions; micropropulsion; micro- and nanosatellite mission analysis; intelligent microinstruments for monitoring space and launch system environments.

Space Science Applications Laboratory: Magnetospheric, auroral and cosmic-ray physics, wave-particle interactions, magnetospheric plasma waves; atmospheric and ionospheric physics, density and composition of the upper atmosphere, remote sensing using atmospheric radiation; solar physics, infrared astronomy, infrared signature analysis; infrared surveillance, imaging and remote sensing; multispectral and hyperspectral sensor development; data analysis and algorithm development; applications of multispectral and hyperspectral imagery to defense, civil space, commercial, and environmental missions; effects of solar activity, magnetic storms and nuclear explosions on the Earth's atmosphere, ionosphere and magnetosphere; effects of electromagnetic and particulate radiations on space systems; space instrumentation, design, fabrication and test; environmental chemistry, trace detection; atmospheric chemical reactions, atmospheric optics, light scattering, state-specific chemical reactions, and radiative signatures of missile plumes.