Chandra monitoring and trends analysis: status and lessons learned

Bradley D. Spitzbart, Scott J. Wolk, and Takashi Isobe

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

ABSTRACT

The Chandra X-Ray Observatory was launched in July, 1999 and has yielded extraordinary scientific results. Behind the scenes, our Monitoring and Trends Analysis (MTA) approach has proven to be a valuable resource in providing telescope diagnostic information and analysis of scientific data to access Observatory performance. We have created and maintain real-time monitoring and long-term trending tools. This paper will update our 2002 SPIE paper on the design of the system and discuss lessons learned.

Keywords: Chandra, trending, real-time systems, automation

1. INTRODUCTION

The core MTA team consists of three people with support from other Science Operations Team (SOT) members, the Flight Operations Team (FOT), and Chandra X-ray Center Data Systems (CXCDS) programmers. The group is charged with providing data, tools, and analysis to access telescope performance as it affects the science quality and efficiency of the Observatory in concert with the engineers, instrument experts, and scientists. This is a unique program to augment the purely health and safety, subsystem-specific approach of the FOT engineering experts.

Section 2 reviews and updates the MTA products and services. We have created and maintain real-time monitoring tools capable of sending e-mail and pager alerts in the event of limit violations or unexpected spacecraft states detected on any of 40 critical parameters during 3 daily contact periods. The real-time telemetry stream is formatted for easy access through a web browser or wireless device. We also are responsible for monitoring radiation data from other satellites (ACE, GOES) at all times. Eighty-two additional values are checked in near real-time including temperature and electronic hard limits and variable instrument configurations versus planned modes. Finally, approximately 1,500 parameters are averaged on 5-minute intervals and databased. This SQL database is used for long-term trending and is accessible to instrument and operations teams. A database of point source properties for use in PSF studies is under development.

Section 3 offers lessons learned. The Chandra MTA system is under constant development to add capability and to respond to changing spacecraft needs. Automation is essential for our small team to provide 24/7 coverage and to maintain routine tasks. We borrow and reuse as much code as possible while developing new scripts, making version control and proper programming a necessary overhead. Data and hardware backup strategies are also considered.

Section 4 offers ideas for future enhancements.

Further author information: (Send correspondence to B.D.S.)
B.D.S.: E-mail: bspitzbart@cfa.harvard.edu, Telephone: 1 617 496 7339
S.J.W.: E-mail: swolk@cfa.harvard.edu, Telephone: 1 617 496 7766
T.I.: E-mail: tisobe@cfa.harvard.edu, Telephone: 1 617 496 7335

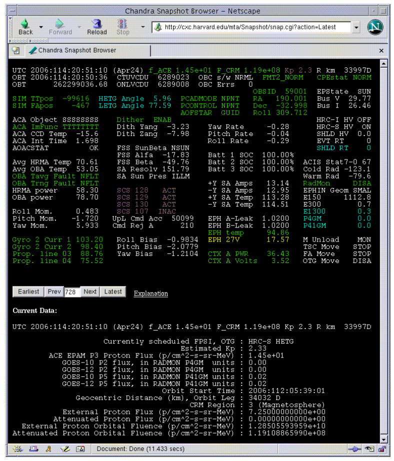

Figure 1. The real-time *Chandra* Snapshot browser.

2. PRODUCTS AND SERVICES

Most MTA products are readily available on webpages, with, of course, some access restriction and password protection in certain areas. The homepage is `http://cxc.harvard.edu/mta/sot.html`. We also produce e-mail and pager alerts and wireless (WML) webpages. At this point, the primary tasks laid out before the mission began have been completed and are now maintained automatically when possible. In addition, the MTA databases are implemented through Phase 2, although some back-filling is still required. Phase 3 will present real science metrics as oppposed to the engineering metrics archived thus far. This includes CTI measurements and emission lines for individual observations.

The majority of MTA time can now be spent on expanding the long-term trending offerings in response to life-limiting expectations and actual anomalies. Finally, we now have more time to create larger, in-depth research projects about once every two months.

2.1. Real-time Data

The *Chandra* Snapshot [*], which provides easy access to the most relevant information from the most recent telemetry. See figure 1. The page is run using a dedicated real-time feed and underlying Perl code to format the ASCII output, limit check, and color code particular items of interest. Alerts are issued on limits violations of the 10 most critical metrics. A cron job checks for new data and updates the values once per minute during real-time contacts. The CGI allows users to review "snapshots" from the past three days and incorporates the latest non-spacecraft data such as ephemeris and radiation environment information. We also provide a text only version for faster access from remote or dial-up connections. A more comprehensive, but less sophisticated real-time display of MSIDs was created for the Flight Operations Team engineers [†]. Known as the State of Health, there are nine pages covering 1011 mnemonics. Pages exist for PCAD (aspect determination), CCDM (command and control), EPS (electrical power supply), Thermal, Propulsion, SIM/OTG (science instrument and transmission grating) Mechanisms, a configuration summary, overall summary (top level), and a safe mode summary for quick assessments of telescope health. The design and implementation is similar to that of the Snapshot, with two dedicated real-time feeds and nine parallel Perl scripts (one for each page) updating every minute during live contacts. There is only limited state checking on the top level and configuration pages and no capability for browsing to earlier epochs.

MTA is also responsible for monitoring the radiation environment at all times. In particular, we gather ACE, GOES-11 (with GOES-8 and GOES-12 as backup), and the Costello Kp index data from solar.sec.noaa.gov.

[*] http://cxc.harvard.edu/cgi-gen/mta/Snap/snap.cgi
[†] http://cxc.harvard.edu/cgi-gen/mta/SOH/soh.html

An ACE P3 (112-187 keV) integrated fluence greater than 3.6e8 particles/cm^2/steradian/MeV within two hours triggers a pager alert. We also now track XMM Radiation Monitor values for each orbit. Although the radiation environment has been mostly benign for the past year as solar minimum sets in, we have started a record of every mission interruption in anticipation of the next solar max and other spacecraft anomalies [‡]. As of April 18, 2006, there are a total of 61 shutdowns caused by high radiation. Of those, 20 were commanded manually from the ground as a precaution, while 41 were triggered autonomously on-board. The number of autonomous trips caused by the EPHIN channels E1300 (2.64 – 6.18 MeV) and P4GM (5.0 – 8.3 MeV) are respectively 30 and 11. None have been caused by the P41GM (41 – 53 MeV) channel alone.

2.2. Stored Data

Chandra stores science and engineering data onboard and downloads it to Earth during each contact. This dumped data is packaged and received from the Ground Operations Team, usually within 2 hours of the end of the contact. The dump is automatically processed by a Perl script. The values are reviewed and compared against expected, as-planned values and operational limits using output products from mission planning and a limits database. When violations are found, e-mails and pages are directed as appropriate. We now monitor 83 mnemonics in this manner, with more being added as required. In addition, transmission grating moves and science instrument moves are captured at this stage and reported on their own web pages. Finally, several ASCII databases are created to supplement the SQL databases used in long-term trending and special reports.

The CXCDS standard data processing pipeline continues to produce, on a daily basis, MTA webpages showing quicklook data on science and CTI observations and limit violations on engineering metrics. The outputs of these pipelines serve as the inputs in creating the SQL databases.

2.3. Long-term Trending

The five minute average and standard deviation is computed for each monitored mnemonic and ingested into an SQL database.[1,2] Currently there are 48 tables containing about 1500 columns, divided by subsystem. An additional 24 tables with 500 columns are slated for future phases. These higher order tables will capture key science data such as emission line detections and point source characteristics.

Extracting values from the MTA databases, we provide plots and statistics for all the monitored metrics (MSIDs). Three iterations are performed on each metric: five minute average values, daily minimum values, and daily maximum values. The system attempts to predict the next six months' behavior by estimating a first and second derivative. The first derivative is simply a first order (linear) fit to the data. The second derivative is a first order fit to the slopes between smoothed data points. Past, current, or predicted future limit violations are highlighted with a green-yellow-red color scheme. Figure 2 shows an example of the trend page output. A subsystem is run each night, to cycle the complete list each week.

The latest strategy in our trending attempts is to define operational envelopes for each metric MSID. We calculate a moving average and note the extrema, minimum and maximum value in each boxcar. The minimum and maximums are then fit with an n–th degree polynomial. In the figure, we have used 1 month boxcars with a 6–th degree polynomial on the left and 10 degree boxes with a 4–th degree fit on the right. The goal is to better describe the actual as-flown data to make the identification of current or future limit violations more apparent and automatic. It may be best to use a linear fit on the past 1 to 2 years for the best predictions. This is work under development.

The flexibility of the web, Perl, and IDL, and the easy access to MTA data allows for the timely creation and presentation of customized studies called for by various teams based on current spacecraft needs or anomalies. Many of these investigations have become new monitoring and trending tasks with the addition of automated scripts and cron jobs. A few examples of the types of customized studies we have done are: tracking ACIS corner pixel, rejected events, and bad pixel trends, timing, position, and voltage requirements for each HETG and LETG insertion and retraction, and PCAD gyro drift rates. Since last reported, studies of the SIM twist, FID light drift, and IRU bias have been added.

[‡] http://cxc.harvard.edu/mta_days/mta_interrupt/time_order.html

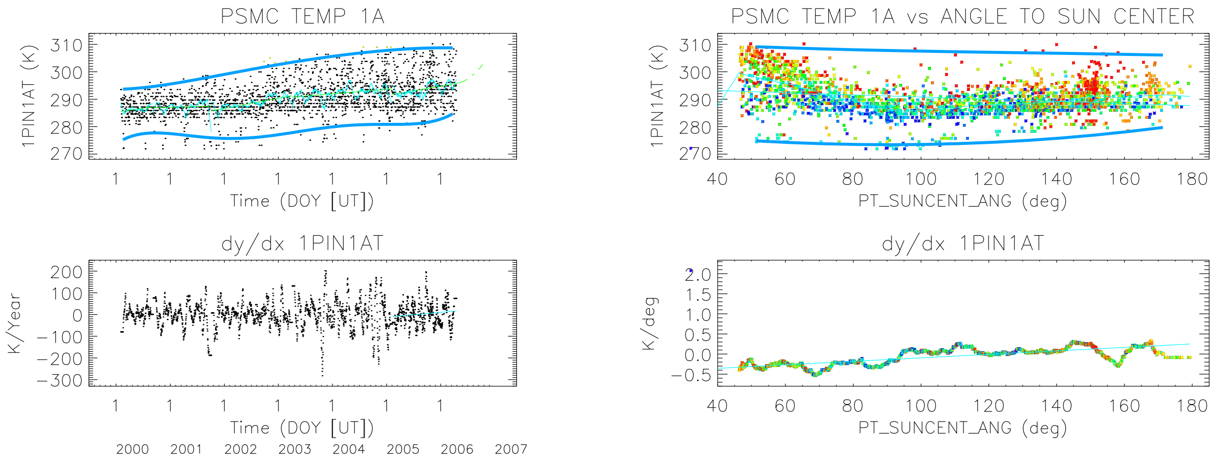

Figure 2. Examples of trending plots from the ACIS subsystem. The IDL code permits plots of MSID versus time (left) or MSID versus another MSID (right). The top panel in each shows the 5-minute averaged data, a linear fit line, and a smoothed curve. The top and bottom blue lines are the envelopes (see sect. 2.3). The bottom panels show the slopes between adjacent points on the smoothed curve and a linear fit, that is the second derivative. The data points in the right-hand plot are color-coded by date in the electronic version. Blue points are the earliest, from 1999; red are the most recent, from 2006.

2.4. Special Reports

The MTA team presents both weekly and monthly reports. The weekly reports focus on limit violations, observations completed, focal plane temperature, and the general health and safety of the vehicle. The monthly report is more in-depth, looking at monthly and year-to-date or past year trends.

In addition we have recently started undertaking in-depth studies of some factors to be presented in memo format. Highlights of a recent study on the HETG energy resolution are shown in Appendix A and on the MTA website.

3. LESSONS LEARNED

A common complication for many groups is dealing with different types of data, input formats, programming languages, and output requirements. Our basic priority is often to accomplish the easiest tasks first over the most critical. This scheme makes sense at the beginning of a mission when most of the laid out objectives have nearly the same criticality. Everything has to be done as soon as possible. At this stage in Chandra's life we may be later than desired in shifting this paradigm to avoid only working on projects that can be done in days and never getting to month-long efforts. Following are our lessons learned for overcoming these problems.

We strive to reduce the human workload as much as possible, redirecting the available resources to new development. Automation is essential. We rely heavily on the UNIX time daemon (cron) to autonomously run jobs at various times of the day, updating data files and web pages and monitoring processing status and telemetry. MTA's crontab consists of nearly 100 periodic tasks. Most of these jobs require some type of watchdog. Some just write out a logfile. Others actively alert an operator if there is a failure. In the case of the real-time system, the most sophisticated watchdogs failover to an independent data flow on another computer. There are several indicators of failure to spark an alert. For example, input data has not updated for a certain amount of time, or has updated in one place but another. We also check that the correct output files or plots are created at the end of each cron instantiation. There are several jobs in which we have not been able to eliminate manual steps. Extended sources must be removed from Science Instrument Background analysis, Hot pixels, cosmic ray hits, and pileup must by considered by a human in total dose calculations, radiation interruptions are entered by hand, and the zero order location is determined by visual inspection in gratings analysis.

At SAO and the OCC, we are blessed with dedicated 24-hour systems personnel to service fileservers, webservers, printers, etc. However, there are several backup techniques employed for data and processing. All of the real-time pages are mirrored on two independent servers. The pages are created from separate data streams processed on two different machines. The software features automatic failovers of the Snapshot including alerts. By default, the backup Snapshot code is responsible for issuing alerts. When data is seen flowing on the primary server, control of alerts is transfered via semaphore to the primary side. At one point we went so far as to set up a Yahoo group as a platform to get messages and data out in an extreme crisis. This has never been used and quickly went inactive. The biggest problem with the many data streams and backup schemes is keeping documentation up-to-date. In some cases only one person may have the intimate knowledge of the inner workings with little current documentation. The use of easily updated webpages helps some. Keeping codes under version control provides built-in documentation.

Most of the MTA scripts had been maintained individually by the code's author. About two years ago we determined to move toward a version control system to better track changes and allow multiple or new employees understand and modify the code. For this task, CVS [§] was chosen because it is open-source, in wide-spread use and simple. Up to this point, only its basic utilities, such as checkout, add, commit, and log have been needed. We use Makefiles and the gmake utility to install upgrades and to ensure that redundant codes receive the same updates with appropriate modifications to path names, IP addresses, etc. We strive to use environment variables and soft links in place of hard-coded paths and filenames. This has been the most time-consuming and frustrating part of the conversion. Usually, our software cycle begins by testing an idea as quickly and simply as possible and evolving the program from there. There is no clear transition from one phase to the other and development goes back and forth from testing to coding. Therefore, it is imperative to begin the testing phase in good programming style, with plenty of comments included. Having never worked in this environment, the transition has not been easy for the team. While all new projects are created under this procedure, some legacy code remains unbound. We would recommend implementing these types of professional programming techniques as early as possible.

Unlike the flight software teams and data processing teams, MTA software is not subject to any review or approval process. This allows for faster, more flexible development. Much of our work is exploratory without defined methods from which to write formal specifications at the onset. We also are allowed to make changes on the fly to suit temporary or contingency plans. This arrangement has generally been agreeable, but some deficiencies are noted. We have not set up a good, robust off-line testing environment. When code changes are made they are often installed directly as the live programs. This has especially been a problem for programs that send e-mail and pager alerts and has resulted in several false alerts.

4. FUTURE ENHANCEMENTS

The MTA effort has been successful in setting up its primary functioning, creating extensive databases, and expanding with the needs of the mission. Many new tasks are under development.

Due to the complicated, dynamic thermal evolution of the spacecraft we seek better HRMA/OBA Thermal visualization. Along these lines, we need a better understanding of the on-board heater duty cycles. There is one known inconsequential stuck-on heater. Some others have been suspected, but exonerated after a painstaking effort.

We already use the Flight Operations Team's Matlab tools to calculate Earth angles and predicted EPHIN temperatures, but seek a more seamless and automated integration. The IFOT system is another asset to be integrated, especially for the comparison and verification of expected versus actual flight behaviors. (See D. Shropshire, this volume, for a full discussion of the Matlab and IFOT tools.)

A huge advance will be to expand and automate our extreme values database. This database holds the maximum and minimum violations throughout the mission. This task includes automating the generation of the daily limit violation table that appears in the weekly report. Likewise, we have an idea to detect bad telemetry by looking for short (1 or 2 sample) limit violations. A couple of corruptions have passed the FOT and checksum verifications during the mission.

[§] http://www.nongnu.org/cvs

There are several processing metrics that could become part of the MTA suite. There is a general sense of the latency required to receive dump data, run the MTA pipes, and populate the databases, but no hard statistics. The OPS team keeps some of this data. We can add run times for our scripts. We currently measure disk space once daily and send an e-mail to the MTA team if capacity is 90% or higher. We can improve the tracking of disk space and CPU usage at various times of the day and night to predict short-term problems, to more cost-efficiently anticipate purchases for upgrades and to guide future missions on needs. We currently keep track of any alerts generated by saving the e-mails to a folder. It would be better to log them automatically in an unencumbered and searchable format.

The entire Science Operations Team is aware of the need for disaster preparedness. The critical MTA software is quite portable and can be run remotely. We will consider and drill for scenarios in which we are forced to quickly change locations, operate remotely, or shutdown and "safe" the system temporarily. This preparation need not apply only to news breaking events. The leased Chandra Operations Control Center space has suffered occasional power outages, mostly planned, due to major construction on the building it is housed in. These are largely a non-event as all operations machines are connected to a generator driven uninterruptible power supply.

5. CONCLUSIONS

Monitoring and Trends Analysis by a Science Operations Team is a novel approach that has proven to be beneficial. The focus of the group on the entire vehicle allows an unique perspective on correlations between subsystems to ensure the highest quality science output possible. In fact, we have evolved through the monitoring and trending tasks and are now concentrating on using the framework for detailed analysis while maintaining our core duties. We hope that our design tips and lessons learned are useful on other missions.

APPENDIX A. E/ΔE ENERGY RESOLUTION TREND

We collected E/ΔE data of several emission lines from all grating observations. The lines used are 0.825 keV, 1.022 keV, and 1.472 keV for ACIS-S HETG, 0.653 keV, 0.825 keV, 1.022 keV, and 1.472 keV for ACIS-S METG, and 0.653 keV, 0.825 keV, and 1.022 keV for HRC-S LETG.

Statistics are computed by a weighted linear least square fit, and a robust fit. To remove the outliers, we first compute an mean and a standard deviation of E/ΔE values, selected data inside of two sigma deviation from the mean. We also dropped data DOM ¡ 100, since the SIM is not at a fixed position yet. The Student's t-probabilities are also computed on these limited data sets. Although some trends seem mildly significant by the weight least square fit, the results by the robust fit or Student's t-probabilities tell us otherwise. Note, the robust fit cannot utilize the errors in E/ΔE value, and also cannot give a formal error for slopes.

Figure 3 shows ACIS-S/HETG E/ΔE trends. The red lines are computed by weighted least square method, and green lines are computed by robust fits as described above. All data points including outliers are plotted, even though they are not used to compute trending lines. Although some of the plots seem to have much less scattering and smaller error bars, plotting ranges of E/ΔE are significantly different from one plot to the other, and scatter is not much different.

We examined whether OBA/HRMA temperature shift affects E/ΔE of observations. The line used in this memo is 1.022 keV line, since all HETG, METG, and LETG observations have the line. The temperature data are collected from MTA data seeker database, and the temperature values used are those of the beginning of the observations. There are several cases we could not find a corresponding temperature data due to gaps in the database. There are also the cases in which recorded temperatures are constant for the entire range. In these cases, we did not plot the relationship (many HRMA temperatures fall into this case).

Although there are a few cases which appear to have a weak correlation between the E/ΔE and the temperature (an example), in general, the OBA/HRMA temperatures do not affect the E/ΔE values, even among those close to the gratings (a statistical significant level is 80% or above for a null hypothesis).

Figure 4 shows ACIS/HETG observations and OBA temperature relations on top. The bottom plot are those with HRMA temperature. The fitted lines are computed by robust linear regression method. Note that the plotting ranges are very different from one plot to the other, and differences in the sizes of the error bars, and the inclines of the slopes are mainly due to the differences in the plotting ranges.

ACKNOWLEDGMENTS

This work is supported by NASA contract NAS8-39073.

REFERENCES

1. R. S. Overbeck *et al.*, "DataSeeker: A Web Interface to Chandra MTA Data," in *Astronomical Data Analysis Software and Systems XI*, D. Durand, D. Bohlender, and T. H. Handley, eds., *ASP Conf. Ser.*, 2002.
2. S. J. Wolk *et al.*, "The Monitoring and Trends Databases for the Chandra X-ray Observatory," in *Astronomical Data Analysis Software and Systems XI*, D. Durand, D. Bohlender, and T. H. Handley, eds., *ASP Conf. Ser.*, 2002.

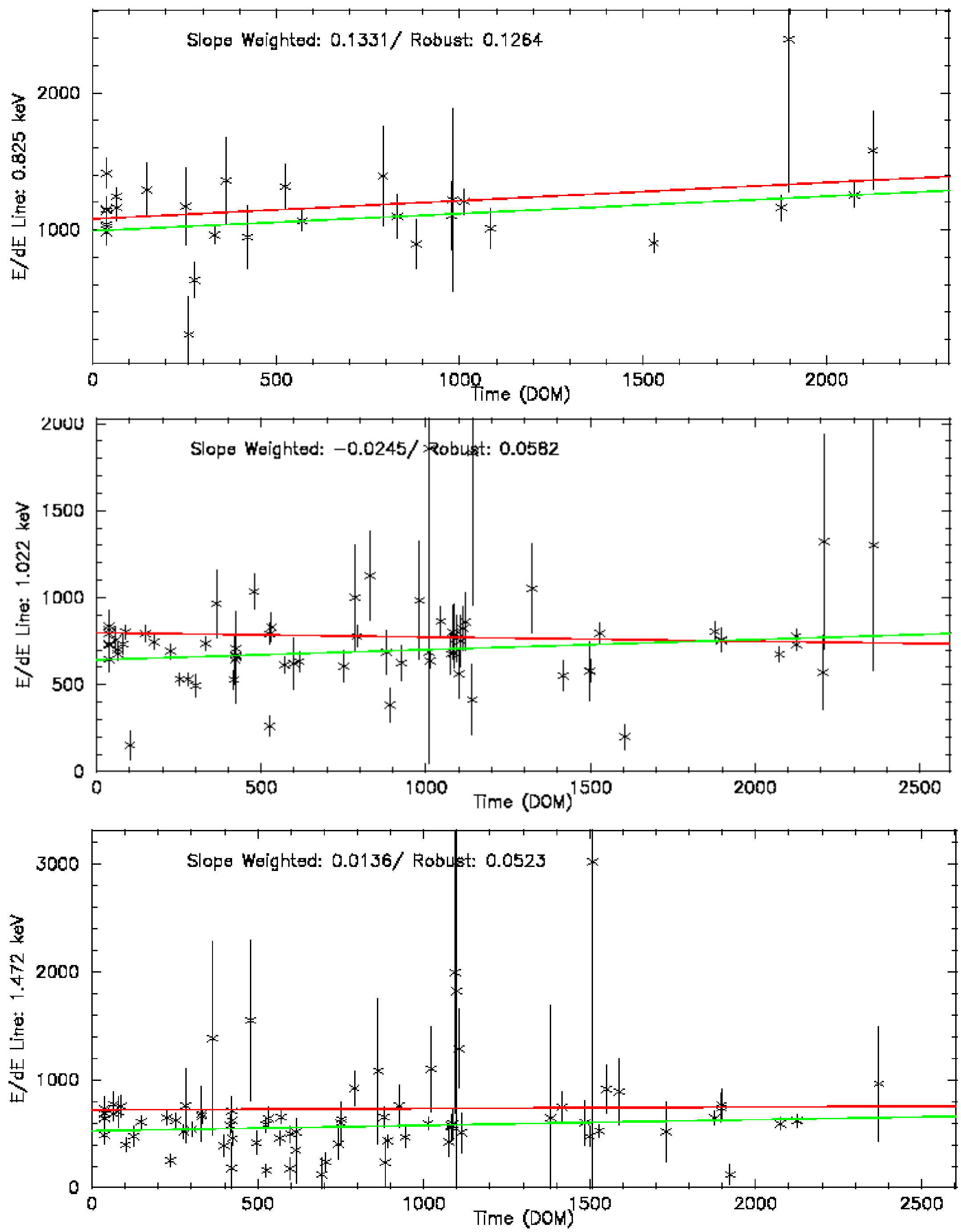

Figure 3. ACIS-S/HETG E/ΔE

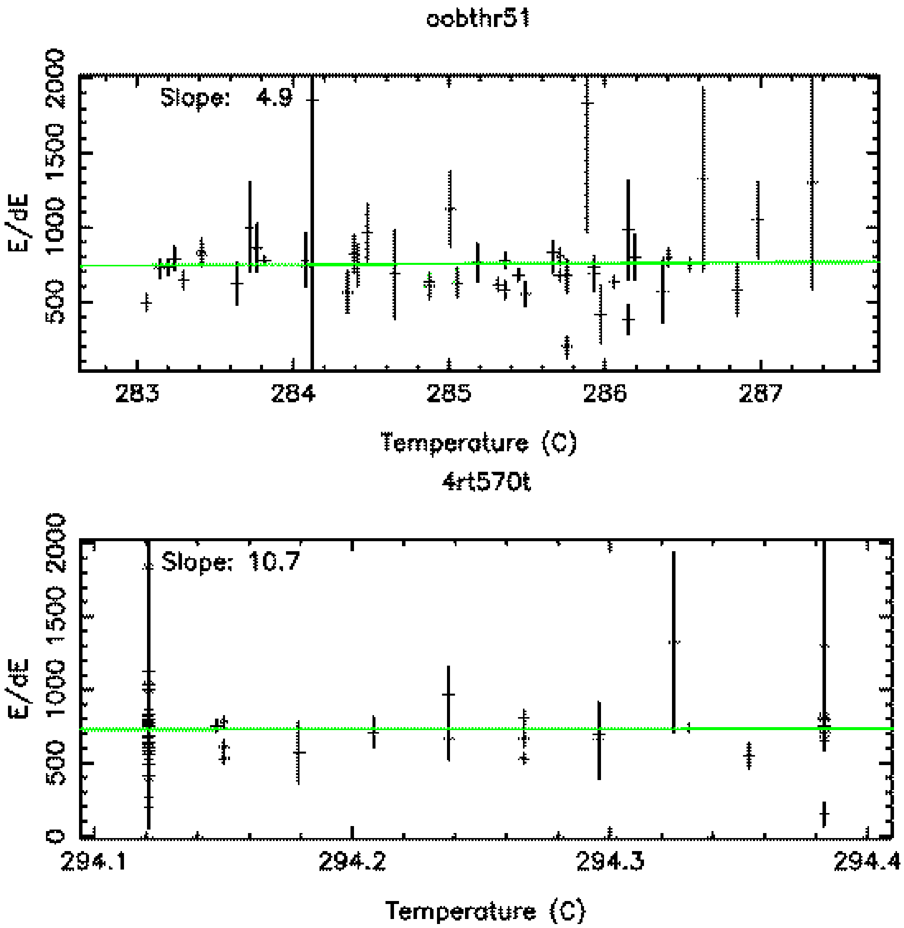

Figure 4. ACIS-S/HETG E/ΔE versus OBA average temperature (top) and versus HRMA average temperature (bottom).