As of 10 June 2016

This GOES-R Level 2+ space weather Algorithm Theoretical Basis Document (ATBD) is preliminary and subject to change as the GOES-R Program prepares for a launch in late 2016.

The NOAA National Centers for Environmental Information (NCEI) is currently developing a demonstration version of the Satellite Product Analysis and Distribution Enterprise System (SPADES) which will host the L2+ algorithms. The operational SPADES will be instantiated by the National Weather Service in FY2017 to serve the needs of the Space Weather Prediction Center (SWPC).

Updated versions of the L2+ ATBDs will periodically be made available and posted to the SPADES website.

Bill Denig
NCEI

http://www.ngdc.noaa.gov/stp/space-weather/online-publications/stp_sii/spades/

NOAA NESDIS CENTER for SATELLITE APPLICATIONS and RESEARCH

GOES-R TIME SERIES DATA AVERAGES ALGORITHM THEORETICAL BASIS DOCUMENT
Version 1.3

TITLE: TIME SERIES DATA AVERAGES ALGORITHM THEORETICAL BASIS DOCUMENT VERSION 1.3

AUTHORS:
Dr. Paul Lotoaniu
Mary Shouldis

APPROVAL SIGNATURES:

____ <Actual Signature Date>
Mary Shouldis
Technical Approver

____ <Actual Signature Date>
Steve Hill
Agency

TIME SERIES DATA AVERAGES ALGORITHM THEORETICAL BASIS DOCUMENT VERSION HISTORY SUMMARY

Version	**Description**	**Revised Sections**	**Date**
1.0	Created by Mary Shouldis CIRES/SWPC	New Document	Oct 15, 2008
1.2	Revised by Paul Loto'aniu	All	Mar, 2009
1.3	Revised by Paul Loto'aniu	All	Dec, 2009

LIST OF FIGURES AND TABLES

LIST OF ACRONYMS

ATBD	Algorithm Theoretical Basis Document
AWG	Algorithm Working Group
CIRES	Cooperative Institute for Research in Environmental Sciences
EUVS	Extreme UltraViolet Sensor
MAG	Magnetometer
SEISS	Space Environment In-Situ Suite
STAR	Center for Satellite Applications and Research
SWPC	Space Weather Prediction Center
XRS	X-ray Sensor

ABSTRACT

This algorithm averages Level 1B high resolution data into averages for the GOES-R magnetometer, energetic particle sensor, x-ray sensor and extreme ultraviolet sensor. Mathematical and pseudo code descriptions are given for the algorithm. Described are assumptions made regarding input data, instrumentation and input spacecraft attitude data and ephemerides. Example test runs using proxy data are shown and compared to results from currently implemented algorithms, with comparisons showing excellent agreement. Pre- and post-launch calibration/validation plans are presented along with current algorithm limitations and possible future improvements.

1.0 INTRODUCTION

1.1 Purpose of This Document

The purpose of this document is to describe the averaging algorithm and its theoretical basis. This document can be used as a reference for implementing the algorithm into software. The averaging algorithm converts 2 Hz (and/or 8 Hz) resolution magnetometer data for all components, total field and inclination into 60 second (~16 MHz) resolution data. This document also provides a description of the methods, mathematical basis, implementation, and testing of the algorithms designed to calculate the one-minute averages of the Level 1b irradiance data from the solar X-Ray Sensor (XRS). The methods for calculating the one-minute and daily averages for the EUV Sensor (EUVS) are also described as well as the one-minute and five-minute SEISS energetic particle averages. Table 1.1 below describes all of the requirements addressed by this algorithm.

Table 1.1 Requirements satisfied by this algorithm

Instrument	Requirement	Time Interval(s) for Averages
Magnetometer (MAG)	MAG.08	1 minute
X-ray Sensor (XRS)*	XRS.04	1 minute
Extreme UltraViolet Sensor (EUVS) **	EUVS.03	1 minute
Space Environment In-Situ Suite (SEISS)	SEISS.16	1 minute
Space Environment In-Situ Suite (SEISS)	SEISS.17	5 minute

*At this time, requirement XRS.03 for the 3-second XRS average is not being addressed because it is assumed the raw L1B will be at 2.7 second cadence which will meet the requirement for 3 second data.

** The requirement for EUVS.03 Daily averages is not covered by this algorithm. It will be addressed separately because it is not a simple average.

Current and previous GOES averaged data are used extensively by the research community (e.g., to validate magnetic field models) and in operations. The algorithm is a simple moving average over the specified span of time of the time series data. When data gaps are present in any specified data period, the decision to average is based on having a minimal number of valid data points. If this minimal number is not met the average is not taken and instead an error fill value is used. Assumptions made regarding input data, instrumentation and input spacecraft attitude are explained along with a discussion on future improvements to the algorithm using weighted averages.

1.2 Who Should Use This Document

The STAR AIT group shall use this document to integrate the algorithm into their collaborative framework environment. It shall also be used by the prime development and implementation contractor to design, develop, test, validate and implement the algorithm into the final operational processing system. In addition, the GOES-R AWG space weather application team shall use this ATBD to verify their operational requirements are being met by the proposed algorithm. They should also use it to understand the strengths and weaknesses of the algorithm as well as its accuracy and applicability.

1.3 Inside Each Section

Section 2.0, OBSERVING SYSTEM OVERVIEW describes the product objectives and requirements, and assumptions made regarding instrument characteristics.

Section 3.0 ALGORITHM DESCRIPTION, describes the theoretical basis of the algorithm, assumptions made for sensory data, input data and ancillary data, algorithm input/output, error estimates and programming, procedural and computational considerations.

Section 4.0 TEST DATA SETS AND OUTPUTS, describes the test data sets used to characterize the performance of the algorithm and the data product quality. Also are described are the results from the algorithm processing on simulated input data.

Section 5.0 PRACTICAL CONSIDERATIONS, discusses issues involving numerical computation, programming and procedures, quality assessment and diagnostics and exception handling.

Section 6.0 ASSUMPTIONS AND LIMITATIONS, discusses algorithm performance with regard to accuracy and the limitations of the algorithm and possible improvements.

Section 7.0 REFERENCES, provides all references mentioned in the ATBD.

1.4 Related Documents

(1) GOES-R Magnetometer PORD (417-R-MAGPORD).

(2) GOES-R Mission Requirements Document, Version 3.0, Feb 2007.

(3) SEISS Performance and Operational Requirements Document (PORD) draft dated November 2004.

1.5 Revision History

Revision Number	Date	Author	Revision Description	Reason for Revision
1.0	Oct 15, 2008	Mary Shouldis	Create document	Create document
1.2	Mar, 2009	Paul Loto'aniu	Revised all sections	Revised all sections
1.3	Dec, 2009	Paul Loto'aniu	Revised all sections	Revised all sections

2.0 OBSERVING SYSTEM OVERVIEW

2.1 Product Generated

The averaging algorithm averages Level 1B high-resolution time series data into averages for specified time spans. The output is average time series data or an indication of an error condition.

The time averaged data will be used extensively by the space physics community and space weather forecaster. For example, observed variations in the earth's magnetoplasma environment of order 1-minute or slower can often be approximated by magnetohydrodynamic (MHD) fluid models, greatly simplifying the plasma physics equations. Furthermore, forecasters and researchers often use 1-minute data to help diagnose general space weather environment.

2.2 Instrument Characteristics

Algorithm development by the Space Weather Algorithm Team assumes instruments meet the performance requirements outlined in the GOES-R MRD and PORD.

For the MAG instrument operational requirements and characteristics are detailed in sections 3.3.6.2 and 3.4.2.7 of the GOES-R Series Mission Requirements Document (MRD) Version 3.0 dated February 2007 and the MAG instrument Performance and Operational Requirements Document (PORD) draft dated November 2004. The accuracy of the algorithm relies on precise instrument calibration and timing. Accurate calibration is critical as errors flow through to the 1-minute averaged product output. For example, stray magnetic fields not accounted for in pre-launch ground testing may create offset errors.

The XRS and EUVS instrument operational requirements and characteristics are detailed in sections 3.3.6.3 and 3.4.2.3 of the GOES-R Series Mission Requirements Document (MRD) Version 3.0 dated February 2007.

The SEISS instrument operational requirements and characteristics are detailed in sections 3.3.6.1 and 3.4.2.5 of the GOES-R Series Mission Requirements Document (MRD) Version 3.0 dated February 2007 and the SEISS instrument Performance and Operational Requirements Document (PORD) draft dated November 2004. The MPS instrument will provide pitch-angle resolved differential fluxes in 9 look directions, from 30 eV- 30 keV for the low energies electrons and protons, 30 keV – 4 MeV for high energy protons and 30 keV – 1

MeV for high energy electrons. The multiple look directions and energy channels are required to correctly represent the measured particle distributions for which the temperature and density moments are being calculated. The accuracy of the algorithm relies on precise instrument calibration and timing.

3.0 ALGORITHM DESCRIPTION

3.1 Algorithm Overview

The time series data averaging algorithm averages Level 1B high resolution data into averages for the specified time span. The averaging method is a simple boxcar method that gathers data over the specified time span and then divides the sum of the data by the total number of data points collected. The output is the averaged time series data. A sliding boxcar average is a simple, well-understood method for calculating the average, and has been used for all previous GOES satellite averages. A window width is selected (such as one minute) and all sensor data measurements falling within that window are given equal statistical weight in the average. That average is assigned to be representative of a single time point within the window. In typical boxcar averaging, the central time point of the window is selected. In previous GOES averaging calculations, the earliest edge of the sliding window has been used. For intervals where the measurements are uniformly distributed in time, this simple averaging technique provides an easily calculated and understood representation of the series data at a particular time.

Any invalid instrument sensor data values are filled in by fill-in error values (e.g., 9999 for the MAG) at Level 1B. All values are accumulated or aggregated whether the vlaues are valid or error fillin values. An erro flag is attached to each values, whith 0 indicating valid value and 1 indicating error value. At averaging time the aggregated values are avergaed based on the percentage of valid values in the aggregate array/vector. If the minimum number of valid values in the aggregate vector is satisified the valid values are averaged, else the averged value is taken as the default error fill-in value.

The averages are calculated, with all values in the time frame receiving equal weighting. For SEISS, the five minute averages are computed from the one-minute averages, with each of the (possible) five values receiving equal weight.

3.2 Processing Outline

The processing outline of the averaging algorithm is summarized in the figure 3.1 below. The current algorithm is implemented as C++ routines with an object oriented emphasis.

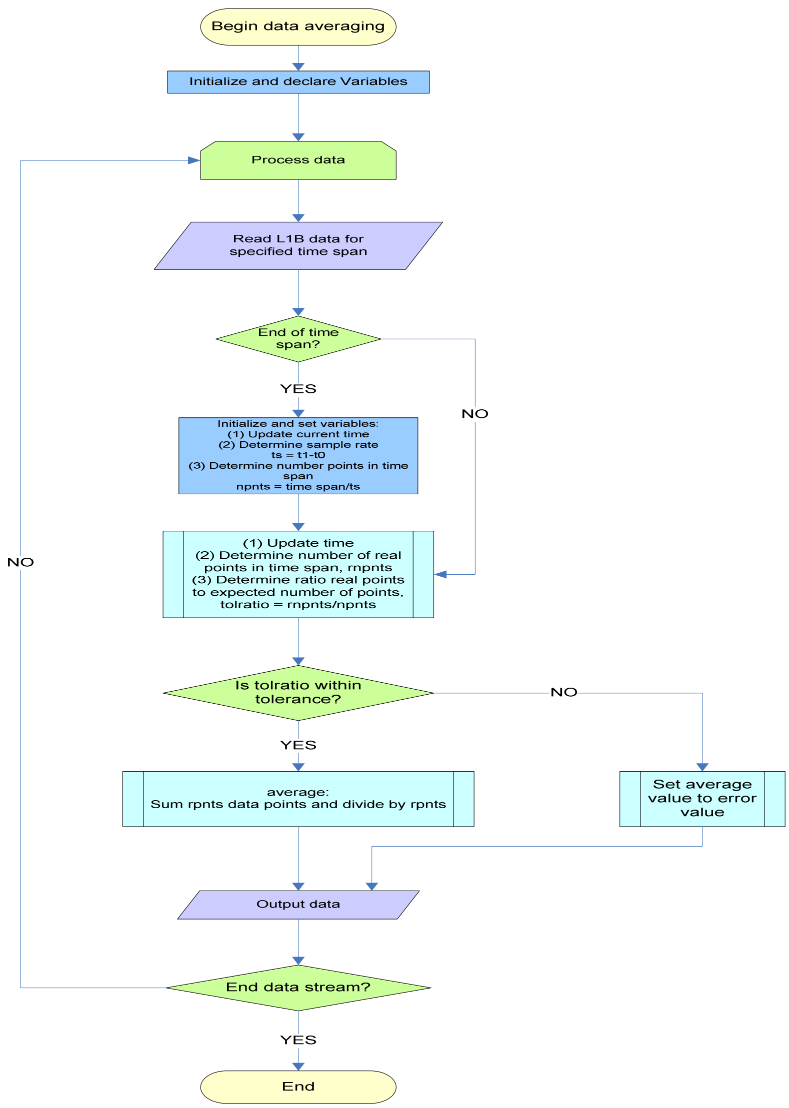

Figure 3.1 Flowchart of the time series data averaging algorithm

3.3 Algorithm Input

3.3.1 Primary Sensor Data

The primary magnetometer sensor data is defined as information that is derived solely from the MAG observations, date/time tags and spacecraft location. The algorithm inputs are

- Sensor data: each 1-D component magnetometer data.
- Date and time: The date and time are converted to meet algorithm input requirement:
  - Time: in milliseconds past previous midnight.
  - Date: in modified Julian date with day accuracy.

For XRS:

- XRS solar soft x-ray irradiance in each of the two band passes at 0.05-0.4 nm and 0.1-0.8 nm.

For EUVS:

- EUVS spectrograph measurements in each of three wavelength ranges.

For SEISS:

- Particle flux data for each energy channel (electrons or protons).

3.3.2 Ancillary Data

There is no ancillary data requirement.

3.3.3 Algorithm Initialization Input Requirement

The averaging object can be initialized using parameter inputs. This will enable the algorithm to be flexible and robust enough to process all of the different time series GOES-R data and to produce the one-minute and five-minute averages. Initialization of the averaging object requires the following inputs:

- Time: starting time in milliseconds
- Sampling rate: Sampling rate in milliseconds
- Averaging length: The averaging length in milliseconds, example 60 seconds averaging would use 60000 millisecond input.
- Maximum time: The time in millisecond where time variables are reset. For example, set to 24 hours (86400000 milliseconds) means time variables are reset at the beginning of every day.

- Error value: The algorithm assumes that errors and data gaps in the sensor measurement data stream are filled with an error value. For example, 9999.
- Minimum good values: The minimum percentage of good data required for averaging to take place, for example, 30 == 30%.

3.4 Theoretical Description

3.4.1 Physics of the Problem

Averaging data minimizes the effect of random measurement errors and allows quality flags to be assigned to the time period based. A sliding boxcar average is a simple, well-understood method for calculating an average and has been used for previous GOES averages.

Data averaging should be a simple process. However, the type of data to be averaged and what the averaged data will be used for can complicate the justifications for using a particular averaging technique over another. For example, GOES MAG data has been used by the space physics community to study plasma waves that may be important to particle acceleration and loss. Frequency analysis of these higher frequency variations in can be adversely affected by aliasing and frequency overlapping as a result of the averaging technique. Previous GOES MAG data has utilized a boxcar averaging technique with the time set to the beginning of the minute. For GOES-R the boxcar method is preserved for consistency. However, future improvements to the averaging algorithm may include weighted averages to minimize aliasing.

3.4.2 Mathematical Description

<u>Boxcar method without overlaps</u>

The boxcar method without overlapping is simply the arithmetic mean, $\bar{x}_{bx}$, over a given time series of length *n,* which can be written as

$$\overline{x_{bx}} = \frac{1}{k} \cdot \sum_{i=1}^{k} x_i$$

where, k is the number of points to be averaged and x_i is the *i*th point. The averaging algorithm is repeated in $x_{k+1} \ldots x_{2k}$ lots until we reach the end of the time series. The time tag for each averaged value is the time at the beginning of the minute corresponding to *i=1*.

The time window width is specified (one minute, five minute or daily) and all measurements falling within that window are given equal statistical weight in the average. That average is assigned to be representative of a single time point within the window. In typical boxcar averaging, the central time point of the window is selected. In previous GOES averaging calculations, the earliest edge of the sliding window has been used. For intervals where the measurements are uniformly distributed in time, this simple averaging technique provides an easily calculated and understood representation of the value being averaged at a particular time.

Any invalid instrument sensor data values are filled in by fill-in error values (e.g., 9999 for the MAG) at Level 1B. All values are accumulated or aggregated whether the vlaues are valid or error fillin values. An erro flag is attached to each values, whith 0 indicating valid value and 1 indicating error value. At averaging time the aggregated values are avergaed based on the percentage of valid values in the aggregate array/vector. If the minimum number of valid values in the aggregate vector is satisified the valid values are averaged, else the averged value is taken as the default error fill-in value.

3.4.3 Algorithm Output

The output from the algorithm is the averaged data values time tagged to the start of each minute.

4.0 TEST DATA SETS AND OUTPUTS

4.1 Simulated/Proxy Input Data Sets

Previous GOES measurements are used as proxy data input for all instruments. No simulated data is used. Example results from each instrument are shown in figures below and discussed in the remaining subsections.

4.2 Output from Simulated/Proxy Inputs Data Sets

MAG

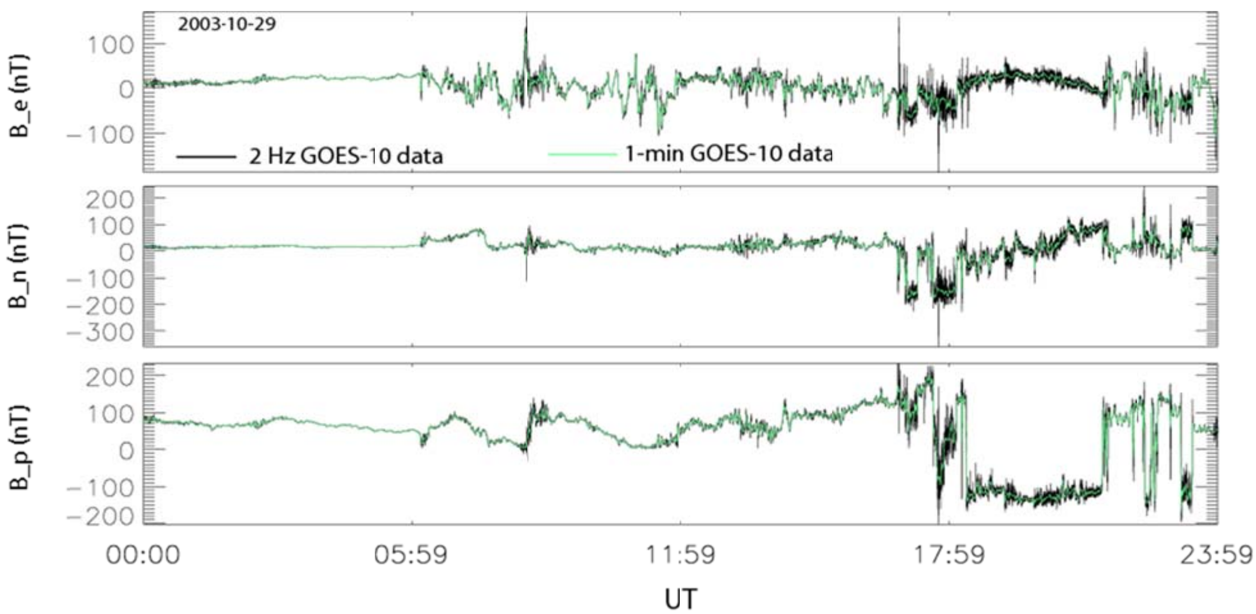

Figure 4.1: Comparison of 2 Hz (0.5 second) data to 1 minute averages from GOES-10 for the three magnetic field components

SEISS

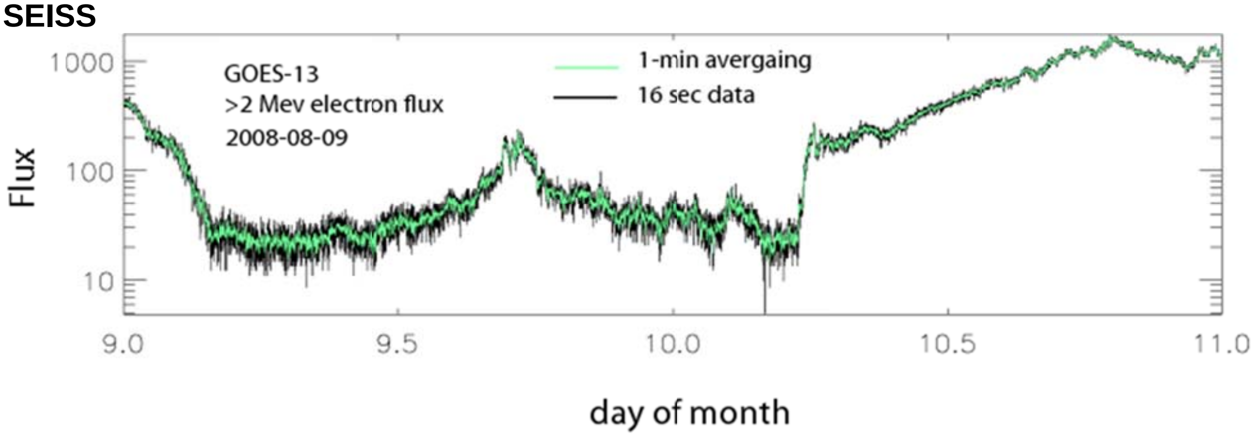

Figure 4.2: Comparison of 16 second >2 MeV electron data to 1 minute averages from GOES-13

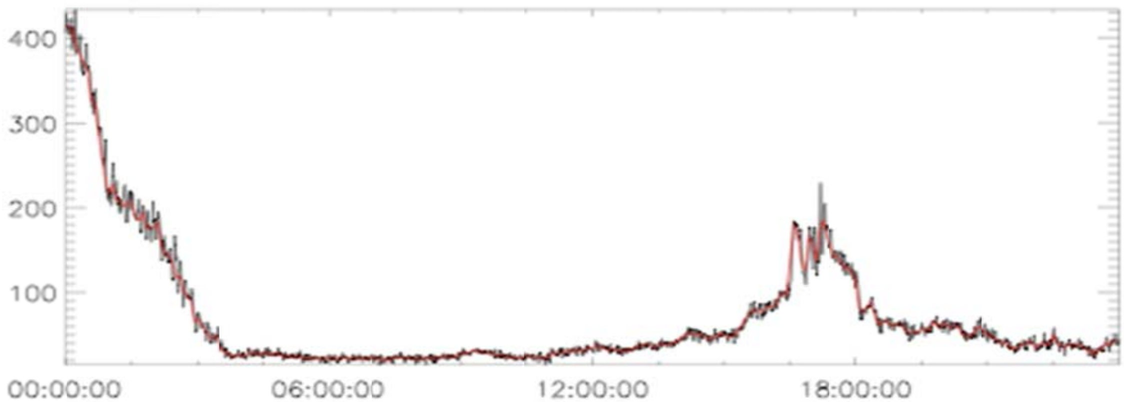

Figure 4.3: Comparison of SEISS 1-min >2 MeV electron data to 5-min averages for 2008-08-09 from GOES-13

EUVS

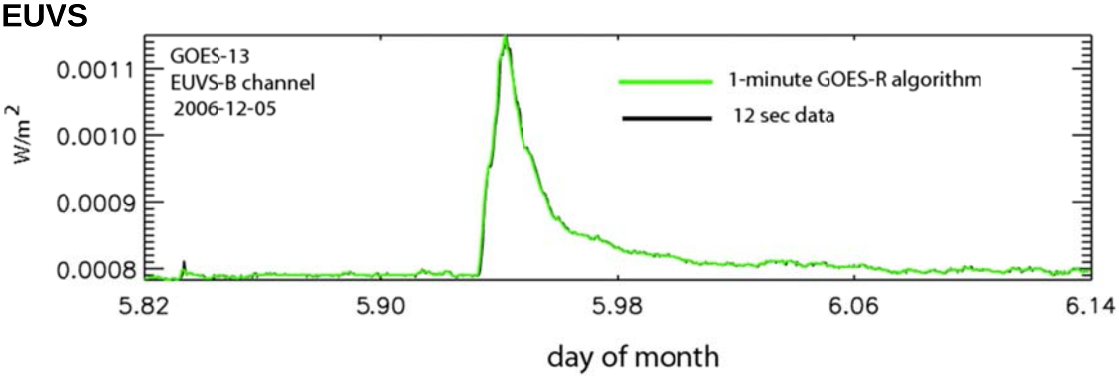

Figure 4.4: Comparison of 27 second data to 1 minute averaged from GOES-13. There are two traces in the plot and comparison is perfect

The exact cadence of the XRS GOES-R sensor is TBD. Future improvements may include scaling the proxy data cadence to that of GOES-R using simple interpolation techniques, then applying the averaging algorithm. For EUVs, this will be the first sensor in the GOES satellite suite to provide continuous space-based EUV measurements. Hence, future improvements includes creation of a proxy data to simulate continous data.

4.2.1 Precisions and Accuracy Estimates

Figures 4.1- 4.4 shows comparisons of GOES-R averaging algorithm output to high resolution data. The agreements are excellent. During highly active periods more data spikes and data gaps appear, hence, error and data gap handling becomes more important. The minimum number of data points required for averaging in the legacy algorithm is unknown. For GOES-R algorithm, the minimum number is a variable that can be changed during each avergaing object initialization.

4.2.2 Error Budget

The minimum number of valid data points required for averaging is a variable that can be set by the user. Currently this number is set at 30%. Hence, 30% of points within the time interval to be averaged must be valid for averaging to be performed else, the error value is returned.

5.0 PRACTICAL CONSIDERATIONS

5.1 Numerical Computation Considerations

There are no computational considerations for the MAG data. However, for the other instruments previous conversions of data from raw counts to physical units and calibration requirements means averaging output will be constrained by the arrival of quality flags and calibration mode information.

5.2 Programming and Procedural Considerations

A separate object must be initiated for each data object to be averaged. Hence, if averaging all three magnetic field components, three averaging objects must be created. After object initialization, the averaging object aggregates streaming sensor data, while returning a *status* flag. For testing purposes, for-loops can be used to simulate the data streaming. The *status* return can take two values:

status = 0: no averaging, still collecting data.
status = 1: avergae data collected, reset counter and update date and time.

Sensor data value errors are assumed to be replaced by an error fill in value, e.g. 9999. The algorithm may need to be updated once actual GOES-R instrument modes and error flags are known. The algorithm will average any constantly sampled dataset where input parameters and variables are known.

5.3 Quality Assessment and Diagnostics

Evaluation of the output as reasonable can be made based on the number of points included in each time interval, as well as the consideration of quality flags, calibration mode information, and dynamic range validity checks.

The following procedures are recommended for diagnosing the performance of the GOES-R averaging algorithm:

- Monitor the percentage of valid data points required for averaging. During very large storms, this minimum percentage may need to be lowered to account for extra data gaps and spikes.
- Routinely compare averaged results to its high resolution equivalent.

5.4 Exception Handling

During the averaging process, a minimum number of consecutive good quality input values are required to return a real average value. Hence, unexpected changes in the time and date as well as data quality must be taken into consideration during implementation of the GOES-R averaging algorithm. Here, we explain the methods integrated as part of the averaging algorithm to account for proxy data quality variations. The averaging algorithm will have to be modified once the GOES-R Level-1B data and quality formats are known.

We use flags to represent different scenarios that may arise in the proxy data. Data quality is represented by a data quality flag, while unexpected changes in the input time and date are symbolized by the time and date flags, respectively.

Sensor Data Quality

The algorithm assumes and applies the following data quality flags for the proxy data:

Table 5.1 Data Flags

BAD	GOOD	error value	fill value
1	0	e.g., 9999	e.g., 9999

If the input value is 9999 a BAD flag is generated, otherwise the GOOD flag is generated. During output, if an average value cannot be generated the default fill value is outputted instead.

Time and Date Quality

The time and date proxy data flags are summarized in the table below, along with functions called as a result of the flag returns.

<table>
<tr><th colspan="3">Table 5.2 Date and Time Flags</th></tr>
<tr><th>Date</th><th>time</th><th>Meaning</th></tr>
<tr><td rowspan="2">0</td><td>0</td><td>Input date equals current date. Input time less than current to time to average. Put back current data value, status=0.</td></tr>
<tr><td>1</td><td>Input date equals current date. Input time is greater than or equal to current time to average. Average data. Status=1, clear data vector and put back current value.</td></tr>
<tr><td>1</td><td></td><td>Input date does not equal current date. Average data. Status=1. Update current date and time.</td></tr>
</table>

5.5 Algorithm Validation

Pre-Launch Validation:
The algorithm has been and improvements will be, validated as described in section 4 using previous GOES data as proxy.

Post-Launch Calibration and Validation:
The algorithm will be calibrated and validated using actual flight level 1B data.

6.0 ASSUMPTIONS AND LIMITATIONS

6.1 Performance

The averaging scheme is simple and equivalent to a boxcar function, with no overlaps. This is consistent with previous GOES averaging. However, the boxcar function introduces high energy frequency side lobes in the spectral domain. Hence, the boxcar method is not ideal for all space physics applications. In particular, wave analysis of the data should take into account this limitation. Error values are assumed to be replaced by error fill in values, e.g. 9999.

6.2 Assumed Sensor Performance

The algorithm assumes sensors will meet specifications and perform within normal operational parameters. When invalid sensor data occurs, it is assumed that these errors are replaced by error fill in values, e.g. 9999. The algorithm will need to be changed once the full range of data flags from the GOES-R sensors, including special mode, calibration mode, etc., are known.

6.3 Pre-Planned Product Improvements

6.3.1 Improvement 1

An improvement to the averaging algorithm will be to use weights in the averaging scheme, which minimize side lobes in the spectral domain. This scheme can be applied to all the space weather instrumentation datasets.

A general weighted value can be defined by

$$y_0(t) = \sum_{k=-M}^{N} W_k y(t + k\Delta t)$$

where, W_k are the weights summed over $-M$ to N data values. In determining the weights the following scheme can be followed:

- Construct theoretical transfer function.
  - What is the required frequency response of the averaged data?
- Estimate W_k such that the actual transfer function approximates theoretical function.
  - Least squares approximation.
  - Chebyshev polynomial approximation.
  - Normalize
  - Determine time-step

The weighted averages will be studied as part of GOES-R risk reduction work and implemented as an improvement in the near future.

7.0 REFERENCES

(1) GOES-R Series Mission Requirements Document (MRD) Version 3.0 dated February 2007.

(2) MAG instrument Performance and Operational Requirements Document (PORD) draft dated November 2004.

(3) SEISS instrument Performance and Operational Requirements Document (PORD) draft dated November 2004