Internal

Review of INST-CAL/ENG Data Archive

2020-07-26 Initial draft by X. Wu

Summary

There is no need to save the input INST-CAL/ENG data; to thin or re-organize the INST-ENG data; and to save the star data. There is no need to save the data in double precision, although it may be too late to change. Archiving in .sav format is allowed, provided that code to convert them to NetCDF is available.

Background

GOES-R extracts certain information from the ABI Level 0 data stream for CWG to monitor instrument calibration. These include the Instrument Calibration (INST-CAL) data, which are ABI measurements of calibration targets; and the Instrument Engineering (INST-ENG) data, which are 52 temperatures of various parts of ABI. The sizes of these data are about 7 TB and 80 GB per ABI per year for INST-CAL and INST-ENG data, respectively. CWG uses these data in two ways. One is to display them as charts on web for users to assess the current status and history. The other is to archive the data for future use. This memo addresses the archive.

The two purposes of converting the input INST-CAL/ENG data into CWG archive are:

1. Re-organize the structure: The INST-CAL/ENG data include all variables in one hour. Future user of the data often need the time series of one or a few variables. If data are archived as they received, one must read a lot of data that is not needed. CWG is concerned with this ineffective I/O, in fact that was why CWG requested for the INST-CAL/ENG data instead of reading them from L0 data.
2. Reduce the size: Nearly 80% of the INST-CAL data are star measurements that CWG may not need (see below). Another 10% are space count, for which statistics are sufficient instead of the 64 individual measurements. CWG archive should be ~10% of the INST-CAL/ENG data.

INST-CAL

CWG reads the INST-CAL data in NetCDF format, separates the data for individual target (ICT, SCT, space, star) and detector, and archives the data in IDL .sav format. In an earlier design, new data are appended to the existing file. It is difficult for this design to add missed or corrected data

in chronological order. One can tag a time stamp to the data, but that nearly doubles the size and is still inconvenient to use. New design solves this problem by putting the data in designated position. To address the concerns that .sav files are limited to IDL applications, code will be available to convert the .sav to NetCDF file.

Current archive saves statistics (min/max/mean/stdv) of Channel 7 star data. There is no use for these data. Channel 2 star data could be useful for calibration if we know the star ID and peak radiance (count is inadequate as these data are calibrated for individual detector). There is no plan to use star data for calibration. If that is to change, we should make that decision before the INST-CAL data are purged.

INST-ENG

The size of INST-ENG data is about 1% of the input INST-CAL files and 10% of the archived INST-CAL files. There is no need to thin further; hourly averages are produced but not meant to replace the data of every second. There is also no need to re-organize, because time series of these data is less likely to be needed and/or computers are sufficiently fast to read all variables.

Input Data

The INST-CAL/ENG data that CWG uses as input are archived in CLASS. CWG does not need to save them.

END