CORONAL DIAGNOSTIC SPECTROMETER
**SOHO**

**CDS SOFTWARE NOTE No. 42**

Version 6 24 March 1999

# A GUIDE FOR THE CDS SCIENTIST

William Thompson
Applied Research Corporation
NASA Goddard Space Flight Center
Laboratory for Astronomy and Solar Physics
Code 682.1
Greenbelt, MD 20771, USA

William.T.Thompson.1@gsfc.nasa.gov

# 1 Overview

This document is for the visiting scientist who expects to come to Goddard to help operate the CDS instrument for a week, and for all other scientists hoping to use CDS data. It is not intended to be a complete guide to CDS. Rather, it is meant to be a kind of overview, a "guide to the guides" if you will, with some (we hope) helpful hints to get you started.

# 2 Designing rasters and studies

CDS science planning is organized by studies and rasters. These concepts require some explanation.

A raster is a basic unit of a scientific observation. At the most basic level, it consists of a series of exposures scanning the solar image across the slit using a combination of scan mirror and slit mechanism movements. The step spacing between exposures is kept constant, as are all other instrument properties such as exposure time. Each raster results in an output data (FITS) file.

A study is a series of rasters strung together to form a complete observation designed to answer a particular scientific question. Examples of studies are shown in the CDS blue book [1].

# 3 Rasters

The tool used to design a raster is MK_RASTER. The use of this program is explained in detail in **CDS Software Note #24**, "CDS Technical Planning Software". Basically, one uses this software to select all the parameters that make up an observation, such as slit, exposure time, etc. It also provides a simulation of what the data would look like for various solar conditions.

The easiest way to make a new raster is by modifying an already existing raster, and the easiest way to do that is by invoking MK_RASTER from within MK_STUDY (Section 4). Normally, one does not call MK_RASTER directly, but calls MK_STUDY instead.

## 3.1 Database access

The `Database access` button allows one to retrieve previously defined rasters. That's a useful way to define a raster, by modifying one that already exists. This button is also the way one saves a raster once it has been defined. It's best to always use the `All` option when storing rasters. This makes sure that all the parts of the raster are stored—the software will take care of ensuring that duplicate entries are not made.

## 3.2 Data source type

One of the buttons on the MK_RASTER menu is labelled `Data source type`. This allows one to estimate what the data would look like for a given solar feature, given the raster settings. It's important to select the correct data source type (ZONE_ID), because MK_RASTER also labels the resulting raster definition with the selection, so that future planners can know what kind of solar feature the raster is best suited for.

The data source types labelled "GIS calibration" are used for when the GIS flat-fielding filaments are to be turned on.

### 3.3 VDS Background

If one is defining an NIS raster, then it is best to leave the VDS background windows enabled. These four small windows in the corners of the CCD are used by the analysis software to measure the CCD readout bias at the time of the exposure. Be aware that if you do decide to disable the VDS background windows, e.g. for faster readout times, then your data will be much harder to calibrate. This is *not* recommended.

### 3.4 VDS data windows

The two NIS spectra are slightly tilted when they fall on the detector. This effect is ignored within MK_RASTER. However, when the commands are generated to implement the plan on the instrument, corrections are made to the defined windows to correct for the tilts. These corrections manifest themselves in one of two ways.

If data windows around selected lines are used, then each window will be positioned separately to correctly account for the location of the spectrum at that wavelength. Thus, spatial positions along the slit should be directly comparable from data window to data window.

However, if full NIS1 and/or NIS2 spectral extraction is requested, then the position and height of the data window for each spectrum will be adjusted to include the full range of the requested portion of the slit at each end of the spectrum. It would then be up to the user of the data to take into account the spectral tilt in the data. If something less than the full height of the slit (120 arcsec) is requested with full NIS spectral extraction, then there is a potential for a problem with the spatial coverage because of the tilts of the spectra. In other words, the spatial coverage at one end of the slit will be different from the other end of the slit. Only when the full slit height is selected does this problem go away.

### 3.5 Compression

There are six forms of compression which are available. These are

**Straight copy:** This is the simplest form of compression, since it is actually no compression at all. Each data value is transmitted down as a 16 bit integer.

`Straight copy` should not be used for NIS data, since the VDS detector only produces 12 bits of data.

**Truncate to 12 bits:** This compression mode is intended specifically for the NIS spectrograph. It achieves a compression factor of 1.33 by only transmitting the lower 12 bits of each data value instead of all 16 bits. Since, for the NIS, the upper 4 bits are all zero anyway, there is no loss of information in using this scheme.

**Variable block word length:** Also known as VBWL, this is the most complex compression scheme on board. It is a lossless compression scheme which intercompares pixels to remove redundancy from the data. The VBWL blocksize can be used to fine-tune the compression process. The total amount of compression achieved will depend on the data, but a rough rule of thumb gives a compression factor of $\sim 2$.

One side effect of using VBWL compression is that it requires a continuous data stream in order to do the decompression. If there are any dropouts in the telemetry, then all the packets of that exposure after the dropout are also lost, because they cannot be decompressed.

**Sum over wavelength:** If one is only interested in spatial information, then one can greatly compress the data by summing over the line profile. Of course, this is a lossy compression—there is no way to reconstruct the profile on the ground.

**Sum over spatial dimension:** This is the converse to the `Sum over wavelength` compression scheme. Instead of summing over the line profile, one sums over the slit to obtain an average line profile. This compression mode is only useful for the NIS spectrograph.

**Sum over windows:** This compression mode is a combination of the `Sum over wavelength` and `Sum over spatial dimension:` compression schemes. The entire data window is summed together into a single number. This compression mode is only useful for the NIS spectrograph.

**Sum over N rows in spatial dimension:** This is a relatively new compression mode that allows one to average over 2 to 16 pixels along the slit to sacrifice spatial resolution for speed. The spatial extent of a single pixel is 1.68 arcseconds and it takes 143 pixels to cover the 4 arcminute extent of the slit. If one instead selects `Sum 2 rows`, then only 72 numbers are returned for every wavelength position instead of 143. If one selects `Sum 4 rows`, then only 36 numbers are sent down in the telemetry, and so forth. The topmost pixel will represent a somewhat smaller area on the detector than the other pixels, because the compression factors do not go evenly into 143. This is adjusted for by boosting the signal in that pixel when the FITS file is created, so that all the pixels can be directly compared. Like the other summing modes, this mode is only useful for the NIS spectrograph.

One side effect that should be noted for all but the first two compression modes, is that they are only efficient if the total number of pixels being read out per exposure is $\leq 64,000$. If you go past this limit, then the compression algorithm slows down tremendously, possibly canceling any gain in the telemetry rate. You can determine the effect each compression mode will have by activating the `Analyse Raster Duration` widget.

## 3.6 Efficient rasters

There are two things to watch out for in designing rasters: rasters which are too big and rasters which are too small.

The program which converts the telemetry into FITS files works on an entire raster at once. A raster which takes several hours to download at the normal telemetry rate may require too much computer memory to convert into a FITS file. In that case, the raster designer should consider breaking the raster up into several pieces.

At the other extreme, it is also possible to have a raster which is too small—especially if the raster is intended to be used over and over again to build up a time series. It takes a certain amount of time to generate a FITS file, regardless of how big it is. Currently, it's taking about a minute for each FITS file, although we hope to get that number down some in the future. Another problem with generating lots of very small FITS files is that it is not a very efficient use of space—each FITS file contains two headers which take up several kilobytes of disk space regardless of how much data is in the file. Also, it makes it more difficult to search for the proper data files.

One trick that can be used to make the raster more efficient while still satisfying the scientific requirements is to use the mirror and slit mechanisms to trick the instrument into taking a time series. It's possible to set the mechanism step size to zero while still having a number of scan positions greater than one. For example, if we wished to take a series of NIS exposures at a given solar position, then we could set the mirror step size to 0 while setting the number of mirror scan positions to 50. We would then get a raster containing 50 exposures at the same solar position.

With the NIS, this trick can only be done through the mirror mechanism. With the GIS, either mechanism or both can be used to generate a time series.

# 4 Studies

Studies are put together from the raster building blocks. The program to do this is MK_STUDY.

## 4.1 Pointing

When we talk about the pointing of a raster, we always refer to the position of the center of a raster relative to the center of the sun.

Pointing in a study can be handled in one of three ways:

1. Directly as part of the study. This should only be used in rare cases. An example of such a study is SYNOP where a series of rasters are made going down the central meridian from the north pole to the south pole.

2. Deferred to the science plan. This will be the most commonly used type of pointing. Rather than embed the pointing directly in the study definition, the user specifies the pointing each time the study is used. If several rasters are defined with deferred pointing, then the same pointing is used for all the rasters.

3. Relative pointing. This is similar to deferred pointing, except that only the pointing of the first raster is deferred. The pointing of the subsequent rasters are defined relative to the first. For example, one could tile four $4 \times 4$ arcmin rasters together to cover an $8 \times 8$ arcmin area, using the following relative pointings.

| X position | Y position |
|---|---|
| 0 | 0 |
| 240 | 0 |
| 240 | 240 |
| 0 | 240 |

## 4.2 Repeating rasters and studies

There are several ways in which one can specify that rasters should be repeated within a study,

One can use a "brute force" method by specifying the same RAS_ID and RAS_VAR several times within the same study. For example, SYNOP calls the same raster nine times. It must be defined this way because each raster uses a different pointing.

Also, the last raster in any study can be repeated. The number of times to repeat the last raster (N_RASTERS1) is given when the study is used—in this way it is similar to deferred pointing above. It is anticipated that this will usually be used when the study consists of a single raster which is to be repeated a TBD number of times, but can used for other studies as well.

A third way to repeat rasters within a study is by specifying it directly within the study definition. However, the last raster in the study will be controlled only from the science plan as explained above. Any repeat count within the study definition for the last raster will be considered only as guidance to the user when forming the plan.

Finally, an entire study can be repeated any number of times (N_REPEAT_S). This is specified when the study is used, in the detailed science plan. One can also specify that the study should be repeated with more than one target (N_POINTINGS). The interaction between N_POINTINGS and N_REPEAT_S requires some explanation.

Suppose that one wished to observe at two separate positions, and toggle back and forth between them ten times. One would then simply set N_REPEAT_S=10 and N_POINTINGS=2 and specify the two target positions. The N_POINTINGS loop is interior to the N_REPEAT_S loop.

Suppose, however, one wished to repeat the study 10 times first at one pointing location and then change pointing and do the same thing again at the second pointing position. Depending on how the study is defined one could do this in one of two ways. If the study consists of a single raster, one could then set the "Repeat last raster" parameter to N_RASTERS1=10 and set N_POINTINGS=2. If, however, the study is more complicated then that, then one has to put two separate entries in the science plan—one for each pointing—each with N_REPEAT_S=10 and N_POINTINGS=1.

### 4.2.1 Effects on onboard tables

In a study consisting of a single raster, and with a single pointing value, the effect of setting the "Repeat last raster" parameter (N_RASTERS1) and the "repeat plan" parameter (N_REPEAT_S) would seem to reduce to the same thing, and in terms of the data returned, they do. However, the ways these two parameters are implemented onboard are quite different.

Setting the "Repeat last raster" parameter simply sets a parameter within the onboard raster table entry for this raster, and thus has no effect on the size of the onboard tables. The same is not true for the "repeat plan" parameter.

Both "repeat plan" and multiple pointings have the same effect in CPT in terms of creating a new series table for each repeat or pointing. Not only does this increase the number of series tables which are used, but also these can be very small which makes inefficient use of the tables. There are several ways the planner can get around these problems.

For simple studies consisting of a single raster, use “Repeat last raster” instead of “repeat plan”.

In many cases, one needs to adjust for solar rotation in order to keep a target within the field of view for a series of observations. One very convenient way to do this is set N_POINTINGS > 1 and give a series of pointings which adjust for the solar rotation. In fact, there's now a button in MK_PLAN which will calculate all the subsequent pointings from the first one. When the normal duration of the study is short, it's best not to depend solely on N_POINTINGS. For example, if a (single raster) study had a duration of 10 minutes, and one wanted to repeatedly observe a target for 3 hours, then one could first set “Repeat last raster” to 6 to pad out to one hour. During that time, the pointing will only drift by 10 arcseconds. Then one could set N_POINTINGS=3 and specify the three pointings to adjust for solar rotation.

### 4.3 Study titles

There are three “title” fields associated with a study. These are:

- The TITLE, a short string describing the study, e.g. “Synoptic study”.
- The acronym (also known as OBS_PROG), a short word, no more than 8 characters, identifying the study, e.g. “SYNOP”.
- The study variation description (SV_DESC), a short string describing the variation on the study, beyond that given by the title, e.g. “Modified to use 50 sec exposure time”.

The title and acronym are tied together, i.e. they have a one-to-one correspondence. However, any study can have a number of variations, each with its own variation description.

Although it is possible to define a new title and acronym for a variation on a study, it's not really desirable to do so. It's better to simply modify the variation description. Otherwise, it makes searching the study definition database all that much harder. There are circumstances when changing the study title and acronym is suitable when one is modifying a study to such an extent that the old title no longer applies. For example, one might change a study originally designed for coronal holes into one for active regions by simply changing the exposure time.

It is necessary to change the study title and acronym if one changes any of the “fundamental” parameters of a study. These are:

- The detector.
- The slit.
- The number of slit and/or mirror positions
- The slit and/or mirror step size.
- The pattern of rasters making up the study (excluding repeat parameters).
- How pointing is handled, fixed/deferred/relative, or the pointing values for fixed or relative pointing.

### 4.4 Defining variations on studies

The easiest way to define a variation on a study is to begin by reading in the current study definition. This is done with the `View/Select Study From DB` button. One can then replace each raster in turn with a different variation on the raster. Then simply change the study variation description and save the study to the database.

### 4.5 Exporting studies

The routine EXPORT_STUDY is used to export a study definition to an ASCII text file. This file can then be transmitted, e.g. as a mail message, to another site and imported into another copy of the database. This allows you, for example, to define a study at your home institution and then send the study definition to be used on the spacecraft.

The route used for study definitions is as follows. They are first sent to the PI group at Rutherford Appleton Laboratories to be reviewed and tested on the engineering model. If they pass that test, then they are sent on to Goddard to be incorporated into the official database so that they can be used on the flight instrument.

Note that every time that the study definition is transferred from one database to another, the ID numbers may change. Thus, a study which may appear as #15 in a user's private database may become #27 by the time it gets incorporated into the official database. All the other ID numbers embedded within the study definition will also change accordingly.

### 4.6 Mixing slits

Because of the need to select a GSET_ID with a GIS observation (see section 5.4), it is not possible to put together GIS rasters with different slits in a single study. The same goes for different zone IDs (section 3.2).

However, there is no such restriction for NIS observations. There, one can mix and match slits if desired. One can also mix together NIS rasters and GIS rasters so long as the GIS rasters all use the same slit and zone ID.

## 5 Planning observations

The program used to plan the observations is called MK_PLAN. Its use is described in **CDS Software Note #7**, "CDS Planning Tool User's Guide".

### 5.1 Writing IAP files

It is the responsibility of the CDS duty scientist to generate and submit Instrument Activity Plan (IAP) files. These files are then used to display the plans of the various instruments in the daily and weekly meetings. To generate the IAP files, simply select `Write IAP` from the `File` menu. It will then bring up a menu to allow you to write the IAP files based on the plan entries, and submit them to the SOHO Science Operations Coordinator (SOC).

The weekly planning meetings are usually held on Fridays, just after the daily planning meeting for that day. During that meeting, the overall plans for the following week are discussed. The scientist who will be on duty that upcoming week must submit by Thursday IAPs with at least the science part filled out for each day of that week (Monday–Sunday).

Each day, the duty scientist makes up a detailed plan for the operations of the following day, starting at 00:00 UT. Another IAP file is then submitted for that day. If the plan is revised, then a new IAP file is submitted.

## 5.2 Reading KAP files

Keyword Activity Plan (KAP) files contain information from all the instrument IAP files, as well as DSN and commanding schedules. These files are generally ingested automatically, but one can also read them in by hand if desired. To read in the latest KAP files, select `Update From Latest KAP` from the `File` menu.

## 5.3 Objects and zones

One thing that can confuse people is the distinction between OBJECT, OBJ_ID, and ZONE_ID:

**OBJECT:** A generic name for a solar feature, or a standard abbreviation for such a feature. There is a SOHO-wide list of such features and their abbreviations. The purpose of OBJECT is to provide something to search on when browsing the catalogs.

**OBJ_ID:** A specific identification for a particular object, to modify the identification given by the generic field OBJECT. For example, if the object was an active region, then OBJ_ID might have the value "AR7954" to give the exact active region number. Other possibilities are heliographic longitude and latitude values, e.g "N45E70" or simple strings such as "S Pole".

**ZONE_ID:** The pointing zone ID, sometimes referred to as the data source type (see Section 3.2). It is used to indicate the expected brightness of the source. In some sense, ZONE_ID plays a similar role as OBJECT. However, its definition is much more restrictive—there are only a few valid ZONE_IDs, while there is a much larger list of valid OBJECTs. For example, coronal streamers, loops, and mass ejections all have separate OBJECT designations, while they would all have the same ZONE_ID. Also, ZONE_ID is a CDS-specific concept, while OBJECT is SOHO-wide.

## 5.4 Selecting a GSET ID

When using the GIS spectrograph, it is necessary to select a GIS setup (GSET) definition. This contains all the parameters needed to set up the GIS detectors for the planned observation. The selected GSET must be consistent with the selected slit and data source type (ZONE_ID)—this will be invisible to the user as the software will only allow GSETs which are consistent with these restrictions to be selected. In general, the most recent GSET should always be chosen.

### 5.5 Running IMAGE_TOOL

IMAGE_TOOL is the program that allows the planner to view recent solar images and select pointing. You can run IMAGE_TOOL from within MK_PLAN, by selecting it from the `Operations` pull-down menu. There are four kinds of solar images that can be selected:

**SOHO SYNOPTIC DATA:** Daily solar images from groundbased telescopes and other solar satellites such as Yohkoh.

**SOHO SUMMARY DATA:** Daily solar images from SOHO instruments.

**SOHO PRIVATE DATA:** Additional data from SOHO instruments. It cannot be emphasized enough that these data are available *strictly for the purposes of planning SOHO observations.* They are not to be used for any scientific data analysis without the permission of the instrument PI.

**PERSONAL DATA:** Any personal directories containing FITS or GIF images.

### 5.6 Running CPT

The Command Preparation Tool (CPT) is used to convert the plan generated by MK_PLAN into commands to send to the instrument. The XCPT program provides a widget-based interface to CPT, and can be invoked from within MK_PLAN, from the `Operations` pull-down menu. CPT should only be run when one is ready to send up the commands—it will get confused if it is run again before the previous commands are sent up. Always coordinate running CPT with the instrument operator.

## 6 Looking at the data

The FITSGEN program converts the telemetry files into FITS format. This program is run by the instrument operator. Observations made during the realtime pass should be available as FITS files soon after the last packet has been received. The data from the overnight periods should be available sometime during the next day.

Simultaneously as the FITS files are written, the data files are cataloged into an IDL database. This catalog can be browsed using the program XCAT. After a day, the data should also be available in an Oracle database which can be browsed via a WWW browser such as Mosaic or Netscape (not yet implemented).

### 6.1 Maintaining the catalog

*(Note: Normally, this is the responsibility of the instrument operator, rather than the planner.)*

Occasionally, the FITSGEN program will be unable to match the science telemetry to the observation plan. The most common cause of this is when the time at which sequences are run falls well out of step with when they were planned to be run. The result is that the science objective and other fields in the catalog will be blank. (If the science objective has been automatically filled in correctly, then all the other parameters will be correct as well.)

The XCAT program can be used to edit the catalog to correct such problems. Simply select the desired entry and press `Edit Catalog Entry`. The following parameters will need to be modified:

SCI_OBJ
SCI_SPEC
OBJECT

For GIS observations, the correct value of the parameter

GSET_ID

will also need to be inserted. Other parameters which may need to be modified, if they were set to anything other than default values, are

PROG_ID
CMP_NO
OBJ_ID
TRACKING
ZONE_ID

In addition, in some cases FITSGEN may not be able to determine the raster ID numbers. In that case, you may need to modify the following parameters:

RAS_ID
RAS_VAR
OBS_SEQ, to match the above, e.g. "R14V1"
EXPTIME (NIS only)

The changes are made when the "Commit" button is pressed. Pressing "Cancel" will cancel all the changes in that editing session.

Another field that may need to be editted is SEQVALID. Set this to 0 if something went wrong, and the data are unusable for some reason.

Proper maintenance of the catalog is important because the information is used to generate the final distribution of the FITS files. Care should always be taken when editing the catalog. In the more complicated cases, e.g. when the raster ID numbers need to be entered, feel free to call on the GSFC support staff for help.

There may also be occasions when it is desirable to add comments to the entries in the catalog. This can also be done from within XCAT. The comments will also be added to the FITS header when the final distribution is generated.

### 6.2 Quicklook and data analysis

The quicklook software allows one to look at the data files themselves. This software is explained in **CDS Software Note #20**, "Quicklook Software User's Manual". Some other software notes relating to data analysis are:

**CDS Software Note #9,** "CDS IDL Data Structure"

**CDS Software Note #13,** "Report on VDS Subsystem Calibration"

**CDS Software Note #21,** "Profile Fitting to CDS/SUMER data"

**CDS Software Note #32,** "A QL Density/Temp. Diagnostic Tool for CDS"

**CDS Software Note #33,** "A Synthetic Spectrum Program for CDS"

**CDS Software Note #34,** "Applying Calibration to VDS Data"

**CDS Software Note #37,** "Density Diagnostic Line Ratios for CDS Using CHIANTI"

**CDS Software Note #39,** "Automatic Line Fitting to CDS Spectra"

**CDS Software Note #41,** "Analyzing CDS Data in IDL: An Observers Guide"

At a more basic level, these software notes describe some lower-level routines designed to help with data analysis and visualization:

**CDS Software Note #5,** "SERTS graphics devices routines"

**CDS Software Note #6,** "SERTS image display routines"

**CDS Software Note #14,** "Time conversion routines"

**CDS Software Note #28,** "Tape Archiving Procedures for the SOHO/CDS"

## 7 More information

For more information, consult the CDS web pages. The CDS home page can be found at

| | |
|---|---|
| Europe: | `http://solg2.bnsc.rl.ac.uk/` |
| U.S.: | `http://orpheus.nascom.nasa.gov/cds/home/` |

Notes about planning can be found at

| | |
|---|---|
| Europe: | `http://solg2.bnsc.rl.ac.uk/cds_images/planner/planners.html` |
| U.S.: | `http://orpheus.nascom.nasa.gov/cds/home/cds_images/planner/planners.html` |

## References

[1] R. A. Harrison. The Coronal Diagnostic Spectrometer for SOHO: Scientific report. version 6.0, 1995.