**GSICS Convention for NetCDF SRF**

 **– Current convention, known issues and proposals for updates**

**Masaya Takahashi** (GDWG Chair, Japan Meteorological Agency)

**Drafted 30 July 2018; updated 4 October 2018**

Instrument Spectral Response Function (SRF) characterizes the sensitivity of instrument’s each spectral band. The SRF files are provided by instrument vendors, satellite operators or user communities in various forms such as csv, xml, MS Excel and flat file. In 2013, GDWG developed a tool to convert such SRF files to netCDF in support of GSICS activities. The user community has used these netCDF SRF files, for example the GIRO (GSICS Implementation of the ROLO model) adopts these files as [one of the inputs](http://gsics.atmos.umd.edu/bin/view/Development/Srf4Giro).

In order to standardize the netCDF SRF as one of *GSICS Deliverables*, the current file naming and its netCDF Convention have been reviewed by GSICS members. Known issues and a proposal to update the netCDF SRF file naming and contents are described in the following sections.

1. **File Naming Convention**

*GSICS Correction* (i.e. inter-calibration products), which is one of the *GSICS Deliverables*, follows [GSICS Filenaming Convention](http://gsics.atmos.umd.edu/bin/view/Development/FilenameConvention). The Convention also follows the rules given in the General File Naming Conventions section of the *W.M.O. Manual on The Global Telecommunication System.* The netCDF SRF generated as one of the GSICS activities was also proposed to follow the Conventions. However, the current file naming is not fully compliant to the WMO Convention due to a wrong use of *DataDesignator* and a lack of date information.

# Example of netCDF SRF file naming (differences are highlighted in red).

Current naming : W\_JP-JMA-MSC,VIS+IR+SRF,Himawari8+AHI\_C\_RJTD.nc

Proposed naming : W\_JP-JMA-MSC,SATCAL+SRF,Himawari8+AHI\_C\_RJTD\_201309--------\_01.nc

VIS+IR+SRF corresponds to *DataDesignator,* which consists of up to three fields delimited by the plus character (+): *DataCategory*, *InternationalDataSubcategory*, and *LocalDataSubcategory*. The first two are mandatory and the last one is optional. *DataCategory* declares the general type of the data while the *InternationalDataSubcategory* provides more specific description of the data. *DataCategories* and *InternationalDataSubcategories* are defined in the *Common Table C-13 of the W.M.O. Manual on Codes*. Each data producer is free to specify its unique *LocalDataSubcategories* for a given pair of *DataCategory* and *InternationalDataSubcategory*. *DataCategory* for all GSICS files is Calibration dataset (satellite), code figure 30. Available GSICS *DataDesignators* are described in Table 1. It should be noted that adding “SRF” to the table (i.e. new *InternationalDataSubcategory*) requires interactions with WMO. This also means the SRFs become one of the GSICS Calibration dataset if “SRF” is listed in the WMO table.

Table 1 Available DataCategory and InternationalDataSubcategory for GSICS (as of July 2018)

|  |  |
| --- | --- |
| **Common Table C-13 Data Category**  | **Common Table C-13 International Data Subcategory**  |
| **Alphanum. Code**  | **Name**  | **Code Figure**  | **Alphanum. Code**  | **Name**  | **Code Figure**  |
| SATCAL  | Calibration dataset (satellite)  | 30  | SUBSET  | Subsetted data  | 0 |
| COLLOC  | Collocated data  | 1  |
| OBC  | On-board calibration data  | 2  |
| BIASM  | Bias Monitoring | 3  |
| NRTC  | Near Real-Time Correction  | 4  |
| RAC  | Re-analysis Correction  | 5  |

“201309--------“ in the example above represents the delivery date of the Himawari-8 AHI SRF to the customer. The date convention is yyyyMMddhhmmss. If a particular date and time stamp field is not specified, it must be replaced by a "-" (minus) character. The example means the SRF information is as of September 2013 with no date information.

“01” just before “.nc” shows *major version number* of [GSICS Versioning](http://gsics.atmos.umd.edu/bin/view/Development/NetcdfConvention#Data_Versioning). This field is optional in the WMO File Naming Convention, but it would be useful in the case of updating/re-evaluating the SRF file.

The use of the product phase (maturity) fields “demo”/ “preop” / “op” is not envisaged as the SRF files are not considered as products. In the case of test data for users, the data version shall be used.

1. **NetCDF Convention**

In a similar way to the file naming, netCDF SRF follows [GSICS netCDF Convention](http://gsics.atmos.umd.edu/bin/view/Development/NetcdfConvention), which is compliant with CF Conventions and Metadata. The following proposal reflects the discussions at 2017 GRWG/GDWG Annual Meeting.

**2.1 Global Attributes**

Different parts between proposed and existing ones are highlighted in red in Table 2. A change of *Conventions* attribute from CF-1.6 to CF-1.8 is proposed (the reason is noted in section 3) even though the CF-1.8 is still in draft. Updating *licence* text is proposed for public use. Two Global Attributes, *processing\_level and* *time\_coverage\_start*, are added in the proposal. *time\_coverage\_end* is not proposed to avoid confusion as this file is actually unknown if new versions are created. However, such information may be useful when new file is created or after the end of instrument’s lifetime, so it might be revisited in future.

Table 2 Global Attributes of netCDF SRF (values are example for Himawari-8 AHI)

|  |  |  |
| --- | --- | --- |
| **NAME**  | **PROPOSED CONTENT**  | **CURRENT CONTENT**  |
| **Conventions**  | CF-1.8 (Note: this is in draft as of July2018) | CF-1.6 |
| **Metadata\_Conventions**  | Unidata Dataset Discovery v1.0 |
| **standard\_name\_vocabulary**  | CF Standard Name Table (v 57, 11 July 2018) | CF Standard Name Table (v19, 22 March 2012) |
| **project**  | Global Space-based Inter-Calibration System http://gsics.wmo.int  |
| **title**  | Himawari-8 AHI Normalized Spectral Response Data  |
| **summary**  | Normalized spectral response functions (SRF) for all channels of the Himawari-8 AHI instrument are stored in this file. The SRF values were normalized by dividing original SRF values by its maximum value. The data came from the official files (see the source global attribute). | Normalized spectral response functions (SRF) for all channels of the Himawari-8 AHI instrument are stored in this file. The data came from the official files (see the source global attribute).  |
| **institution**  | JMA |
| **licence**  | Normalized spectral response functions (SRF) delivered as a GSICS deliverable is generated in accordance with the GSICS principles and practices. Normalized SRF files are public and may be used and redistributed freely. Neither the data creator, nor the data publisher, nor any of their employees or contractors, makes any warranty, express or implied, including warranties of merchantability and fitness for a particular purpose, or assumes any legal liability for the accuracy, completeness, or usefulness, of this information. | This file was produced in support of GSICS activities. Any publication using this file should acknowledge both GSICS and the data's relevant organization. Neither the data creator, nor the data publisher, nor any of their employees or contractors, makes any warranty, express or implied, including warranties of merchantability and fitness for a particular purpose, or assumes any legal liability for the accuracy, completeness, or usefulness, of this information. |
| **creator\_name**  | JMA |
| **creator\_email**  | jma-msc-contact@ml.kishou.go.jp  |
| **creator\_url**  | https://www.data.jma.go.jp/mscweb/data/monitoring/calibration.html |
| **platform**  | Himawari-8 |
| **instrument**  | AHI  |
| **wmo\_satellite\_code**  | 173 |
| **wmo\_satellite\_instrument\_code**  | 297 |
| **source**  | AHI-08\_SpectralResponsivity.zip (MD5:144e1f26d6fd8f6edce5cc6b1a6f6ad9) |
| **date\_created**  | 2018-07-25T09:08:20Z |
| **date\_modified**  | 2018-07-25T09:08:20Z |
| **history**  | e.g. 2018-07-25T09:08:20Z srf2nc.py v3.0.13  |
| **id**  | W\_JP-JMA-MSC,SATCAL+SRF+VISIR,Himawari8+AHI\_C\_RJTD\_201309--------\_01.nc | W\_JP-JMA-MSC,VIS+IR+SRF,Himawari8+AHI\_C\_RJTD.nc |
| **naming\_authority**  | jp.go.kishou.mscweb |
| **publisher\_name**  | JMA |
| **publisher\_email**  | jma-msc-contact@ml.kishou.go.jp |
| **publisher\_url**  | https://www.data.jma.go.jp/mscweb/data/monitoring/calibration.html |
| **processing\_level** | v1.0.0 | Not exist |
| **time\_coverage\_start** | 2013-09-01T00:00:00Z | Not exist |

**2.2 Dimensions**

There is no change between current and proposed netCDF SRFs.

Table 3 Dimensions of netCDF SRF (values are example for Himawari-8 AHI)

|  |  |  |
| --- | --- | --- |
| **NAME** | **VALUE** | **DESCRIPTION** |
| **channel** | 16 | # of channels included in the netCDF file |
| **sample** | 5449 | Max # of SRF samples among channels in the netCDF file |

* 1. **Variables**

Table 4 represents variables in the proposed and the current netCDF SRF, respectively. Different parts are highlighted in red. Several attributes exist in the Convention (e.g. *\_FillValue*, *valid\_min* and *valild\_max*) are not shown in the table. *origin* is a flag which represents original sample domain (i.e. “1” / “2” denotes that original SRF was measured in wavelength / wavenumber unit). Two CF Standard Names related to wavelength are proposed to be added (standard\_name: radiation\_wavenumber is not listed in the CF Standard Names table). Optional variables on uncertainties of wavelength, wavenumber, and SRF are proposed by GSICS members for the future missions (e.g. MTG/FCI). An extension of the convention to other spectral domain such as microwave may be discussed in future.

Table 4 Variables of the netCDF SRF.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **NAME (Dimension)**  | **LONG\_NAME**  | **STANDARD\_NAME**  | **UNITS**  | **TYPE**  |
| **channel** (channel)  | nominal channel central wavelength  | sensor\_band\_central\_radiation\_wavelength | um | double  |
| **channel\_id** (channel)  | channel identifier  | sensor\_band\_identifier  | N/A | string  |
| **origin** (channel)  | original sample domain  | N/A | N/A | ubyte  |
| **wavelength** (sample, channel)  | wavelength  | radiation\_wavelength  | um  | double  |
| **wavenumber** (sample, channel)  | wavenumber  | N/A  | cm-1  | double  |
| **srf** (sample, channel)  | normalized spectral response  | N/A  |  1 | double  |
| [optional] **wavelength\_unc** (sample, channel)  | uncertainty of wavelength  | N/A | um  | double  |
| [optional] **wavenumber\_unc** (sample, channel) | uncertainty of wavenumber  | N/A  | cm-1  | double  |
| [optional] **srf\_unc** (sample, channel)  | uncertainty of normalized spectral response  | N/A  |  1 | double  |

1. **Discussions**

**3.1 Adoption of WMO File Naming Convention**

The current netCDF SRF File Naming was decided in the lunar calibration activity to use the files in GIRO (GSICS Implementation of the ROLO model). Following GSICS File Naming Convention (which also follows WMO GTS File Naming Convention) was proposed and adopted in 2014 because the files were thought to be one of *GSICS Deliverables* in future. However, it is not clear whether all the deliverables should follow the WMO Convention.

WMO GTS File Naming Convention is very helpful if the dataset should be discoverable via WMO-relevant systems such as WIS (WMO Information System). It is also reasonable for following such standard format to facilitate a clear identification of the dataset’s (e.g. GSICS Correction) origin and content, and is important in the infrastructure maintenance (e.g. file uploading).

However, above-mentioned points might not be the cases for netCDF SRF because SRF is the information to characterize the instrument response (not a product nor an observation).The SRF files are not critical to the infrastructure because they are typically not frequently updated. As previously mentioned in section 1, we should also be careful about adding “SRF” to *DataDesignator* because this means the netCDF SRFs would be considered as one of GSICS “Calibration dataset”. In addition, there are other issues for some agencies. For example, “RJTD” of the example is an international four-letter location indicator (called “CCCC”) of the station or centre originating or compiling the bulletin published in [WMO-No. 9, Volume C1, Catalogue of Meteorological Bulletins](http://www.wmo.int/pages/prog/www/ois/Operational_Information/VolumeC1/CCCC_en.pdf). It might be difficult for some agencies to find proper “CCCC”. One advantage of adopting the WMO Convention is a clear identification that the files are generated under the GSICS framework.

For these reasons, there are 3 options.

* Option#1: keeping the current convention until we have real requirements to change.
* Option#2: adopting the proposal which follows WMO GTS File Naming Convention.
* Option#3: defining a new convention which is not relevant to WMO Convention.

Option#1 would be reasonable if there are no urgent or real requirements. Option#2 would have no impacts on the current applications, but more discussions are needed to propose a guideline which kinds of GSICS deliverables to follow GSICS and WMO File Naming Convention. Option#3 (e.g. himawari8\_ahi\_srf\_v01.nc) means that the SRF file naming does not follow the WMO Convention any more. This may have impacts on GIRO because GIRO gets satellite and instrument names from the file name, by checking the positions of comma (,), plus (+), and underscore (\_).

**3.2 Required updates for CF-Compliant dataset**

Global Attribute *Conventions*=”CF-1.6” which is contained in the current netCDF SRF means that the file follows CF-1.6. However, the file is not CF-compliant because string and unsigned-byte data types are not supported in the past/latest CF-Conventions. Adding these data types is now discussed in the CF community (discussions on [string](https://github.com/cf-convention/cf-conventions/issues/139) and [unsigned byte](https://cf-trac.llnl.gov/trac/ticket/166)), so the netCDF SRF would be fully CF-compliant in future by updating *Conventions* attribute from CF-1.6 to 1.8.

Alternatively, the file could be CF-compliant by changing data types (i.e. string to char, unsigned byte to byte), but some existing applications which use the netCDF SRF (e.g. GIRO) may be affected by such a change. Keeping the existing convention as it is and updating to the proposed one in future would be reasonable from the viewpoint of avoiding extra resources for this issue.

**3.3 Discussions with GSICS members**

Option#1 was selected through a discussion on gsics-dev mailing list: <https://groups.google.com/forum/#!topic/gsics-dev/8SbHfALbufo>. Option#2 could be discussed in future.

**Acknowledgement:**

Special thanks to Sebastien Wagner, Peter Miu, Claude Ledez, Tim Patterson, and Tim Hewison of EUMETSAT, members of GSICS Data management Working Group, and Aleksandar Jelenak of HDF-Group for their kind and productive comments and suggestions on CF and WMO Conventions.