DIWG Recommendations Compliance Scorecard for Mature NetCDF-4 Products

DIWG Recommendation (Below) \ Product Name (Across)	OMI_MINDS_NO2d (Version 1.1)	NSIDC-0630 (Version 2.0)	MODIS_A-JPL-L2P-v2019.0 (Version 2019.0)	tos_Omon_E3SM
RFC-028 2.1 - maximize HDF5 /netCDF-4 interoperability via API accessibility	Yes (file is in netCDF-4 implementation based on HDF5)	Yes (file is in netCDF-4 implementation based on HDF5)	Yes (file is in netCDF-4 implementation based on HDF5)	Yes (file is in netCDF-4 implementation based on HDF5)
RFC-028 2.2 - include basic CF attributes - "Conventions"	Yes	Yes	Yes	Yes
RFC-028 2.2 - include basic CF attributes - "units"	Yes	Yes	Yes	Yes
RFC-028 2.2 - include basic CF attributes - "long_name"	Yes	Yes	Yes	Yes
RFC-028 2.2 - include basic CF attributes - "standard_name"	Yes (for all 3 coordinate variables)	Yes (7 variables)	Yes (5 variables)	Yes (4 variables)
RFC-028 2.2 - include basic CF attributes - "_FillValue"	Yes	Yes	Yes	Yes
RFC-028 2.2 - include basic CF attributes - "valid_min/valid_max" or "valid_range"	Yes	Yes	Yes	No
RFC-028 2.2 - include basic CF attributes - "scale_factor" and "add_offset"	N/A (no scaling done)	Yes	Yes	N/A (no scaling done)
RFC-028 2.2 - include basic CF attributes - "coordinates"	Yes	No (ACDD content_coverage_type attribute used instead)	Yes	No
RFC-028 2.2 - include basic CF attributes - use UDUNITS values for "units"	Yes (5 variables have "units" attribute attached with **l egal** UDUNITS values), No (3 ColumnAmountNO2 variables have "units" attribute attached with value "molec/cm2", which is UDUNITS **illegal**) and N/A (3 bounds variables and crs variable **do not** have and **s hould not** have "units" attribute attached)	Yes	Yes	Yes
RFC-028 2.3 - use CF "bounds" attribute	Yes (for all 3 coordinate varaiables)	No (bounds variables not used, and so bounds attribute not used)	No (bounds variables not used, and so bounds attribute not used)	Yes
RFC-028 2.4 - verify CF compliance	Yes (no errors and no warnings found in check vs. CF-1.8 via IOOS Compliance Checker and NCAS Compliance Checker)	Uncertain (CF-1.9 implemented in product, while IOOS Compliance Checker and NCAS Compliance Checker can only check up to CF-1.8)	Mostly CF-complaint (one error and many warnings found in check vs. CF-1.7 via IO OS Compliance Checker and NCAS Compliance Checker)	CF-compliant (READING CF Compliance Checker gives "INFORMATION message" regarding "history" attribute for "tos" variable)
RFC-028 2.5 - distinguish clearly between HDF and netCDF packing conventions	N/A (no packing done)	Yes (packing done, with both packing_convention and packing_convention_description attributes included)	No (packing done, but with neither packing_convention nor packing_convention_de scription attribute included)	N/A (no packing done)
RFC-028 2.6 - when to employ packing attributes	N/A (no packing done)	Yes (3 variables of type short and ushort are packed, with packing attributes included) and No (1 variable of type short has packing attributes scale_factor = 1.0 and add_offset = 0.0, and so is **not** actually packed)	Yes (8 variables of type byte or short are packed, with packing attributes included)	N/A (no packing done)
RFC-028 2.7 - mapping between ACDD and ISO	N/A (no ISO metadata included)	N/A (no ISO metadata included)	N/A (no ISO metadata included)	N/A (no ISO metadata included)
RFC-028 2.8 - make HDF5 files netCDF4-compatible and CF-compliant within groups	N/A (no groups included)	N/A (no groups included)	N/A (no groups included)	N/A (no groups included)
RFC-028 2.9 - include time dimension in grid structured data	Yes	Yes	Yes	Yes

RFC-028 2.10 - order dimensions to facilitate readability in grid structure datasets	Yes	Yes	Uncertain (warnings generated regarding dimension order in check vs. CF-1.7 via IO OS Compliance Checker)	Yes
RFC-028 2.11 - consider "balanced" chunking for 3-D variables in grid structures	? (chunk size = 1 x 60 x 60)	N/A (no chunking done)	? (chunk size = 1 x 2030 x 1354)	N/A (no chunking done)
RFC-028 2.12 - include "datum" attributes for data in grid structures	No ("crs" variable used instead)	No ("crs" variable used instead)	No	No (but may not be needed for model data)
RFC-036 3.1 - character set for user-defined groups, variables and attributes	Yes (CF character set used)	CF character set used	CF character set used	CF character set used
RFC-036 3.2 - consistent units for variables across one data collection	Yes	Yes	Yes	Yes
RFC-036 3.3 - use "units" attribute only for variables with physical units	No (units = 1 from CF used for two variables)	No (units = count from UDUNITS used for 1 variable)	Yes	Yes and No (main variables have units, and bounds variables' units equal their coordinates' units)
RFC-036 3.4 - include time coordinate in swath structured data	N/A (grid not swath)	N/A (grid not swath)	Yes	N/A (grid not swath)
RFC-036 3.5 - keep coordinate values in coordinate variables	Yes (identified as coordinate variables via CF coordinates attribute)	Yes (identified as coordinate variable via ACDD coverage_content_type attribute)	Yes (identified as coordinate variable via ACDD coverage_content_type attribute)	Yes (CF coordinates attribute not used because all coordinates have same variable name as dimension name, i.e., all are standard NUG 1D coordinates)
RFC-036 3.6 - include georeference information with geospatial coordinates	Yes ("crs" variable used)	Yes ("crs" variable used)	Not Sure ("WGS84" is mentioned in a comment, but that's all)	No (not needed since CF defaults to spherical Earth and lat/lon coordinates are provided)
RFC-036 3.7 - Not-a-Number (NaN) value	Yes (no NaN values in product)	Yes (no NaN values in product)	Yes (no NaN values in product)	Yes (no NaN values in product)
RFC-036 3.8 - standardize file extensions for HDF /netCDF files	Yes (uses .nc filename extension)	Yes (uses .nc filename extension)	Yes (uses .nc filename extension)	Yes (uses .nc filename extension)
RFC-036 3.9 - ensure granule's filename uniqueness across different dataset releases	Yes	Yes	No	Yes
RFC-036 3.10 - adopt semantically rich dataset release identifiers	Yes	Yes	Yes	Yes
RFC-036 3.11 - date-time information in granule filenames	No (have YYYYmMMDD and YYYYmMMDDthhmmss in filenames instead of YYYYMMDD and YYYYMMDDThhmmssZ)	No (have YYYYDDD in filenames instead of YYYYMMDD)	No (have YYYYMMDDhhmmss in filenames instead of YYYYMMDDThhmmssZ)	No (have YYYYMM in filenames instead of YYYY-MM)
attach the CF flag_values or flag_masks attributes along with the CF flag_meanings attribute to each flag variable	N/A (no flag variables)	Yes (one flag variable with flag_values and flag_meanings attributes attached)		
avoid use of the missing_value attribute	Yes	Yes		
define the projection ellipsoid to match the reference datum	N/A (no projection ellipsoid)	Yes		
document missing granules for instruments that acquire data on a regular basis	No (missing granule information should go on the product home page)	No (missing granule information would go on the product home page)	No (missing granule information would go on the product home page)	N/A (may be no missing granules for model data)
include only one variable per GeoTIFF file	N/A (not GeoTiff)	N/A (not GeoTiff)	N/A (not GeoTiff)	N/A (not GeoTiff)

indicate in CRS metadata the order of elements in horizontal coordinate pairs	N/A (no coordinate pairs)	N/A (no coordinate pairs)		
make a variable's valid data range consistent within each product release	Yes	Yes		
make a variable's valid data range useful	Yes	Yes	Yes	No (no valid data range included. All simulation data are physically valid in this dataset, though ranges could be specified)
use a number outside of the valid data range for a variable's fill value	Yes	Yes	Yes	Yes (though no valid data range included)
use DOIs for referencing documentation	Yes	Yes (URL form of DOI used in "citation", "metadata_link" and "source" file-level attributes, and non-URL form of DOI used in "id" and "references" file-level attributes). Non-DOI referencing is also included, because not all relevant references have DOIs, and to provide additional information above-and-beyond the DOI.	No	Yes and No (esm_paper_reference attribute uses URL form of DOI, but references attribute does not include a DOI)
use double precision when archiving time in seconds since a specific epoch	Yes	Yes		
use only officially supported compression filters on netCDF-4 and netCDF4-compatible HDF5 data	Yes	Yes	Yes	Yes
use the null character as the fill value for string data and metadata	Yes (the only dynamic strings that could go bad are too simple to mess up)	Yes (hopefully the values of the dynamic strings time_coverage_start and time_coverage_end will never be messed up)		
Total Score	**29.5 / 34 = 0.868** **(with 10 N/A)**	**31 / 38 = 0.816** **(with 6 N/A)**		