

Style Guide for Variable Titles in CMIP6

Martin Jukes.

Version 0.1, 21st December 2018.

1. Introduction

The variable titles (which appear in the NetCDF files as values for the “long_name” attribute) should be short phrases that describe the variable at a level of detail suitable for the title of a figure or table. This document sets out some styling rules.

The phrase should be capitalised as described in section 2, and there should be no full-stop.

Some aspects of this guide are not cleanly implemented in 01.00.29.

2. Capitalisation

The long names should be capitalised as titles, following "standard" practise.

1. The first word is capitalised (i.e. the first letter is capital, the rest lower case);
2. Every subsequent word is also capitalised, apart from exceptions listed in (3) and in section 2;
3. The following words stay in lower case, except at the start of the title:
 - Articles: a, an, the
 - Coordinating Conjunctions: and, but, or, for, nor, etc.
 - Prepositions (fewer than five letters): on, at, to, from, by, etc.

See [this](#) for more discussion. Note that some words, such as “as” and “that” can take different roles depending on the structure of the sentence so it is not possible to automate everything easily. But, in the context of CMIP variables, there is a limited range of usages.

The words "neither", "either" are conjunctions when used together with "nor", "or", but otherwise an adjective, pronoun or adverb (e.g. "Neither of the proposals is acceptable"). The usage in CMIP6 is always as a conjunction. Some style guides suggest that "neither" should be capitalised because of its length, but here we opt to keep consistency across the pair "..neither .. nor ..", as in "Percentage of Grid Cell That Is Land but neither Vegetation Covered nor Bare Soil".

The conjunction “where” is used in phrases such as "Surface Temperature Where Land or Sea Ice", in which it is interpreted as a subordinating conjunction. The distinction between coordinating and subordinating conjunctions is unclear in the context of phrases which do not have a clear grammatical structure.

In CMIP6, the following words are kept in lower case:

1. a
2. as at by in of on or to
3. and any but for non nor not out per the
4. into onto over than with
5. either neither degree

3. Special words

Abbreviations and technical terms may have their own capitalisation rules. Exceptions are listed in Annex 1.

4. Hyphens

The general rule is that hyphenation is probably unnecessary if the phrase is used as a noun, but probably necessary if used as an adjective, e.g "Sea-ice Temperature". We should avoid excessive hyphenation of terms, but use it consistently.

The following terms are always used as adjectives, and so consistently hyphenated: "clear-sky", "land-use", "land-cover", "all-sky", "above-ground", "below-ground" and "sea-salt" (depricated, as "seasalt" is preferred).

Others may be used either with or without: "sea-surface", "cloud-top", "grid-cell", "sea-ice", "land-ice".

Denoting interfaces: 'ice-ocean', 'snow-ice'

Other hyphenated terms: '2D-field', 'Aerosol-Free', 'Aqueous-phase', 'Eliassen-Palm', 'Gas-phase', 'Half-Levels', 'Layer-integrated', 'Near-surface', 'Non-orographic', 'Non-woody', 'Oxygen-17', 'Oxygen-18', 'Snow-to-Ice', 'X-component', 'Y-component', 'non-living', 'non-renewable', 'carbon-13', 'carbon-14'.

Rules on capitalizing are a little unclear, but several sources recommend capitalising both words if the 2nd is a noun or adjective (e.g. [here](#) and [here](#)). Following this rule, all the CMIP6 hyphenated phrases have both components hyphenated except "Run-off" and "Snow-to-Ice".

5. Issues

- NO: Nitrous oxide, but could perhaps be a negative ... need to check context;
- as, that: potential for different capitalization in some grammatical constructions;

Oxygen and Carbon Isotopes

Currently using Oxygen-17, Oxygen-18, 13Carbon, 14Carbon: should harmonise this.

Numbers and Units

100m 10hPa 10m 1m 2000m 2m 300m 550nm 700m 865nm ... and others.

A number of titles contain a number of terms with scientific units, e.g. "m", "hPa". Some style guides argue that these should be preceded by a space (e.g. [NIST guide for use of SI units](#), page vi). But [Cambridge University guide](#) advises using no space, as in "10m". This looks better in phrases such as Maximum Wind Speed of Gust at 10m, so we follow the latter approach. We will still leave a space if a unit is spelled out, as in "50 day", "20 degree Celsius Isotherm" (degree is lower case as a unit).

Annex 1: word lists

The python package autocorrect is used, which has a list in autocorrect.word.KNOWN_WORDS.

Names etc

13C 13CO2 14C 14CO2 17O 18O 1H 2H C2H2 C2H6 C3 C3H6 C3H8 C4 CFC11 CFC113 CFC12
CH3COCH3 CH4 CO CO2 D H2 HCFC22 HCl HNO3 HO2 N2 N2O NH3 NH4 NHx NO NO2
NO3 NOx NOy O O1D O2 O3 OH Ox SF6 SO2 SO4 T UGRID X XY Y

Substances, named diagnostics and places

- 4XCO2
- CWD
- DMS: dimethyl sulfide
- NMVOC: non methane volatile organic compounds
- NPP: Net primary production
- PBL: Planetary boundary layer
- pH: acidity
- PM1.0, PM2.5, PM10: particulate matter of less than x micron radius
- TEM: transformed Eulerian mean
- TOA: top of atmosphere

Instruments

- CALIPSO
- CFAD
- CloudSat
- ISCCP
- MISR
- MODIS
- PARASOL

Proper names

- Eliassen
- Vaisala

Technical words not recognised by autocorrect

- Dianeutral
- Diazotrophs
- Downwelling
- Epineutral
- Landuse
- Longwave
- Meltpond
- Mesozooplankton
- Methanogenesis
- Methanotrophy
- Microzooplankton
- Needleleaf
- Picophytoplankton
- Streamfunction
- Submesoscale
- Thermosteric