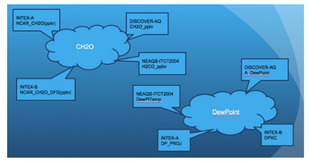
IN32C-0399 - The Use of Atmospheric Composition Variable Standard Names in Airborne and Field Data Products

Morgan L. Silverman[1,2], Gao Chen[2], Michael Shook[2], Megan Buzanowicz[1,2], Rebecca Hornbrook[3], Luke Ziemba[2], John Nowak[2]

[1]Science Systems and Applications Inc., [2]NASA Langley Research Center, [3]National Center for Atmospheric Research

Introduction

The number of variables measured during airborne field campaigns has increased more than tenfold over the last thirty years. With this increase in measurements, the complexity for distributed active archive centers (DAACs) to distribute the data and for data users to search for and find measurements of interest has also increased. Part of this complexity arises from the unique variable names in suborbital atmospheric composition field studies. With limited guidelines related to variable naming, variable names and structures can vary significantly, even for the same type of variable. It is common for instrument scientists to use their intended measurable quantity as the data variable name. This can make it difficult for users to locate and interact with a particular variable across multiple data sets.



One effective solution to this problem, identified by the Earth Science Data System (ESDS) ICARTT Refresh Working Group [1], was to introduce variable standard names that can be used as tags for each data variable. This allows similar measurements (e.g., dew point) to be categorized and located across field campaigns, regardless of what variable name the instrument scientist has used.

From this the atmospheric composition variable standard names were developed with the goal to use Findable, Accessible, Interoperable, and Reusable (FAIR) principles [2] and provide context for all users, while remaining connected to those in the subject area. These standard names have been successfully implemented in FIREX-AQ, CAMP^{2}EX, ACTIVATE, and DCOTSS field campaigns.

Known Issues in Current Naming Systems

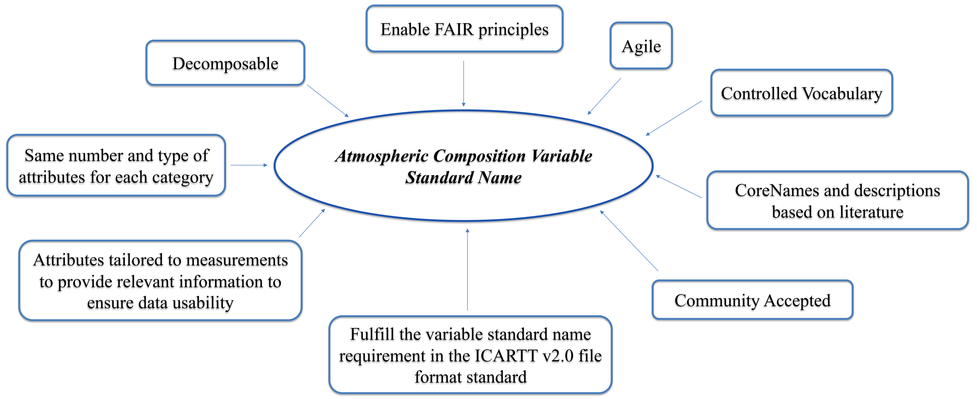
Current variable standard naming systems such as the Network Common Data Form (netCDF) Climate and Forecast (CF) Metadata Convention [3] and the Scientific Variables Ontology (SVO) [4] do not adequately cover the needs of NASA suborbital atmospheric composition field studies, which often include several research grade measurements.

Three main issues exist that hinder the implementation of FAIR principles

1. Measurand coverage and description - Standard names only cover a quarter to a third of the atmospheric composition relevant measurements taken in suborbital field campaigns
2. Structure and vocabulary - Construction of CF names are governed by a set of guidelines that allow for the addition of qualifiers to a base standard name using underscores, allowing an unknown range of possibilities
3. Usability - Some CF standard names are ambiguous and could refer to several different measurements, whereas for other measurements more than one CF standard name could be used.

One naming system that uses a similar approach to the Atmospheric Composition Variable Standard Names is the Aerosol, Clouds and Trace Gases Research Infrastructure (ACTRIS) [5]. Both follow the I-ADOPT [6] approach using a decomposed structure and controlled vocabulary to support the FAIR principles.

Approach



Atmospheric Composition Variable Standard Names

Atmospheric Composition Variable Standard Name = *MeasurementCategory*_***CoreName***_*AcquisitionMethod*_*DescriptiveAttributes*

- Decomposable structure allows a wider range of standard names without specifying each individually thereby creating a very long list; enhances reusability and interoperability.
- Each component (MeasurementCategory, CoreName, AcquisitionMethod, DescriptiveAttributes) is governed by a list of controlled vocabulary, commonly used in literature to provide interoperability.
- The Practical Guide for Atmospheric Composition Variable Standard Names and Controlled Vocabulary can be located at https://www-air.larc.nasa.gov/missions/etc/AtmosphericCompositionVariableStandardNames.pdf
- An RFC is currently in review by the NASA ESDIS Standards Coordination Office (ESCO)

MeasurementCategory

- Can be used with the CoreName to conduct a broad search (Findable)
- 13 categories (Gas, AerComp, AerMP, AerOpt, CldComp, CldMicro, CldMacro, CldOpt, Met, GasJValue, AquJValue, Platform, and Rad) based on shared properties and/or types of measurements
- Type and/or number of attributes within each category are constant and tailored to each type of measurement (e.g., aerosol particle optical property vs. aerosol particle composition) or medium (e.g., trace gas vs. aerosol particle.

MeasurementCategory	Description	Number of Descriptive Attributes	DescriptiveAttributes
Gas	Abundance, relative abundance, or properties of specific trace gas compounds or a group of trace gases measured or reported as one lumped quantity	2	MeasurementSpecificity, Reporting
AerMP (Aerosol Microphysical Properties)	Aerosol microphysical properties of particles not segregated by chemical composition, e.g., abundance, relative abundance, size, and size distribution	4	MeasurementRH, SizingTechnique, SizeRange, Reporting
AerComp (Aerosol Particle Composition)	Aerosol chemical (or composition) properties (including as a function of particle size), i.e., abundance or relative abundance of the chemical components, bulk chemical characteristics, and mixing state	3	SizingTechnique, SizeRange, Reporting
AerOpt (Aerosol Optical Properties)	Intensive and extensive optical properties and optical hygroscopicity of all sampled aerosol particles or a subgroup of aerosol particles	4	MeasurementRH, WL, SizeRange, Reporting

CoreName

- CoreNames and respective definitions can be used for general search when combined with the MeasurementCategory (Findable)
- Many CoreNames are based on common terminology used in atmospheric composition literature
  - When this is not possible (e.g., name too long) the definitions define and explain terminology

Example Trace Gas CoreNames and Definitions

CoreName	Definition	Chemical Formula	CAS Number	Specificity
H2	Hydrogen	H2	1333-74-0	S
iButeneAnd1Butene	Sum of Isobutene and 1-Butene	C4H8	N/A	M
Z13Pentadiene	(Z)-1,3-Pentadiene	C5H8	1574-41-0	S

Example Aerosol Microphysical CoreNames and Definitions

CoreName	Definition
NumConc	Number concentration of aerosol particles
NonVolatileNumConc	Non-volatile number concentration of aerosol particles
CCN	Cloud condensation nuclei number concentration
MassSizeDist	Mass Size Distribution i.e., mass concentration expressed as a function of aerosol particle size.

Examples

Gas_CoreName_AcquisitionMethod_MeasurementSpecificity_Reporting
In-situ measurement of CO2 gas reported in molar fraction with respect to dry air: Gas_CO2_InSitu_S_DMF

AerMP_CoreName_AcquisitionMethod_MeasurementRH_SizingTechnique_SizeRange_Reporting
In-situ measurement of aerosol particle number size distribution reported at reduced relative humidity derived from an aerodynamic sizing technique for coarse-mode aerosols at standard temperature and pressure:
AerMP_NumSizeDist_InSitu_RHd_Aerodynamic_Coarse_STP

AerComp_CoreName_AcquisitionMethod_SizingTechnique_SizeRange_Reporting
Example of an in-situ measurement of bulk sea salt particles reported in number fraction:
AerComp_Seasalt_InSitu_None_Bulk_NumFrac

AcquisitionMethod

- Sampling technique of the measurement

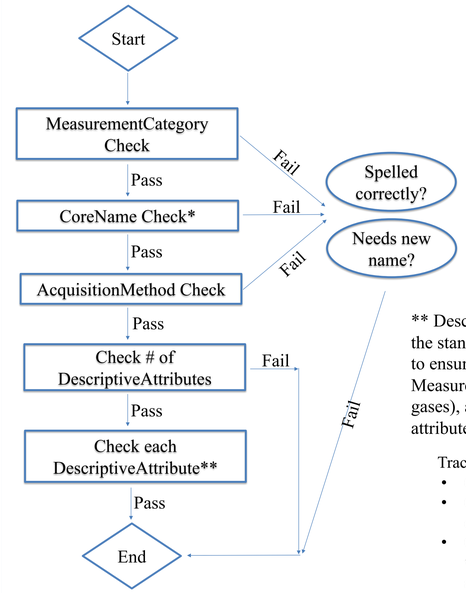
AcquisitionMethod	Description
InSitu	Sampling in close proximity of the instrument or the sampling platform
VertCol	Measurement of a remotely sensed vertically integrated column, where the column measured is nominally perpendicular to the earth's surface
SlantCol	Measurement of a remotely sensed vertically integrated column, where the column measured is not nominally perpendicular to the earth's surface (e.g. the instrument is sun-tracking)
Profile	Measurement of vertically resolved profile

DescriptiveAttributes

- Provide measurement and/or data reporting information relevant for data use and faceted data search (Findable and Reusable)
- The number and types of descriptive attributes are dependent on the MeasurementCategory, but constant within each MeasurementCategory
- Each descriptive attribute field, e.g., WL, SizeRange, etc., has a controlled list of terminology (Interoperable)
- Overlapping descriptive attribute categories, e.g., Reporting, have different lists of controlled terms for each category, reflecting the practice that one measurement can be reported in different ways for different applications
- Some MeasurementCategories, e.g., Meteorology, do not have DescriptiveAttribute. In this case, "None" is used as the value.
- The DescriptiveAttributes for trace gas and aerosols are listed in the Measurement Category table

Machine Interoperability

The Atmospheric Composition Variable Standard Names are designed to be machine actionable. An example of the feasibility of an automated checker to minimize user mistakes is shown below. This is a critical step to ensure the correct use of the standard names. Using a similar approach, one can parse the standard names and convert and/or translate them to CF-like variable names using the controlled vocabulary tables.



* Four possibilities exist when checking the CoreName: correct, belongs to a different category, spelled incorrectly, or needs a new CoreName

** DescriptiveAttributes relate to all components of the standard name. Each attribute must be checked to ensure compatibility with the MeasurementCategory, CoreName (for trace gases), and AcquisitionMethod when reporting attributes are used.

Trace gas example

- Check the number of DescriptiveAttributes (2)
- Check that the Specificity attribute (S/M) corresponds to the CoreName
- Check the Reporting attribute is compatible with the acquisition method, e.g., if VertCol, attribute must be CNDAMB, which is Column integrated number density reported at ambient temperature and pressure

References

[1] Northup, E., Chen, G., Aikin, K., Webster, C., & Batchelor, D. (September 2016). *ICARTT Format Enhancements based on ESDSWG Recommendations* (Publication No. ESDS-RFC-029). https://www.earthdata.nasa.gov/s3fs-public/imported/ESDS-RFC-029.pdf

[2] Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R.,…Goble, C. (2016) The FAIR Guiding Principles for scientific data management and stewardship, *Scientific Data*, 3. https://doi.org/10.1038/sdata.2016.18

[3] Eaton, B., Gregory, J., Drach, B., Taylor, K., Hankin, S., Blower, J., Caron, J., Signell, R., Bentley, P., Rappa, G., Höck, H., Pamment, A., Juckes, M., Raspaud, M., Horne, R., Whiteaker, T., Blodgett, D., Zender, C., Lee, D.,…Herlédan, S. (2022, August 31) *NetCDF Climate and Forecast (CF) Metadata Conventions*, https://cfconventions.org/Data/cf-conventions/cf-conventions-1.10/cf-conventions.html

[4] *Scientific Variables Ontology (SVO)*, 2019, https://scientificvariablesontology.org/svo/index.html

[5] *ACTRIS Vocabulary*. (n.d.). ACTRIS. https://vocabulary.actris.nilu.no/skosmos/actris_vocab/en/

[6] Magagna, B., Moncoiffé, G., Devaraju, A., Stoica, M., Schindler, S., Pamment, A. (2022). InteroperAble Descriptions of Observable Property Terminologies (IADOPT) WG Outputs and Recommendations. Research Data Alliance. https://doi.org/10.15497/RDA00071

[7] Practical Guide for Atmospheric Composition Variable Standard Names and Controlled Vocabulary, May 2022, https://www-air.larc.nasa.gov/missions/etc/AtmosphericCompositionVariableStandardNames.pdf