Problems

William A Cooper

2022-06-06

Contents

Some problems encountered when trying to update the ProcessingAlgorithms document to include all variables appearing in netCDF files. For the record, the procedure used was this:

1. Search for variable names using constructVarTable.R. I ran this on my local computer and also on barolo, searching both /scr/raf_data and /scr/raf/Prod_Data.

2. This script also tried to do some preliminary classification into types, including 'pa' (in ProcessingAlgorithms), 'st' (status or housekeeping), 'a' (needs addition to ProcessingAlgorithms), 'sp' (special use and therefore not needed in ProcessingAlgorithms).

3. The resulting data.frame, with unique variable names, was saved as DVALL.Rdata.

4. I then merged the resulting 'a' variables into Appendix B and tried to link to appropriate descriptions, which had to be added in most cases.

5. In the process, I also did lots of editing to remove remnant HTML code and make the document more consistent with RMarkdown. This mostly involved changing the reference and link style, recording tables to use kable instead of HTML, and removing lots of <p>-</p> tags while ending paragraphs with two spaces. I left some HTML, however, notably the box-paragraph styles used for algorithm boxes. There are probably some remnants I missed.

6. I also edited Appendix C to update how the document should be edited.

7. I then combined the list of problems below and the former "punch list", making it a separate document stored in the www directory so it could be edited easily and referenced apart from the TN. I updated this list where possible.

The following lists some unresolved problems:

1. Some variables appeared only in SCMS files so they have been omitted as too old to be needed here. The archived data frame DV includes a few 'SCMS' types for this reason.

2. I have been unable to learn how BAROCOR is used. It has values around 1013 for some cases but unrealistic ranges for others. It probably refers to the baro-loop correction applied to ALT to obtain pressure altitude, or maybe it is like an altimeter setting. I have omitted it.

3. DOFOR has conflicting definitions, in some cases (OTREC, WCR-TEST) DOF Reject Count, in others (SPICULE, WECAN) 2DC Shadow OR Count. I decided to reclassify it as 'st' and omit it from Appendix B.

4. BLONA appeared in the search but not in the EOL search so I omitted it and replaced BLONGA iN the PA document.

5. DMS appeared only in VOCALS so I classified it as 'sp' and omitted it from Appendix B.

6. GGALTC is defined is "Corrected GPS Altitude" but I can't find where there is any difference from GGALT. It lists PALT among dependies, which is worrisome. I have omitted it pending clarification of the algorithm and its usefulness.

7. GGALTF (GPS Altitude in Feet, Corrected) occurs only in ORCAS, I think. I have omitted it.

8. I decided to omit GGALTSD and re-type it is 'st'. Also, GGLATSD and GGLONSD.

9. GGLATB and GGLONB were only in FRAPPE so I omitted them as evidently the same info as GGLAT/GGLON.

10. GGSF_A was only in CONTRAST so I omitted it.

11. GLATF_A and GLONF_A (Avionics GPS Latitude and Longitude Fine) were only in CONTRAST so I omitted them.

12. Many projects (including NOMADSS, WECAN, WCR-TEST, WINTER) had 'Hybrid' INS variables. My memory from an early test is that I wasn't convinced these were an improvement. I have omitted them pending revisit of my earlier study, which I don't remember well. I have left their status as 'a' though.

13. GVSPD_A appears only in CONTRAST, among projects since about 2010, but I included it anyway as an alternate to GGVSPD.

14. H2O_PIC{2311, 2402, CU} have been widely used, but there was no description of them in ProcessingAlgorithms. I added them (and CO_PIC) to the description for CO2/CH4_PIC. Maybe these entries should have their own paragraphs?

15. I omitted HNO3 and XHNO3 (FRAPPE and NOMADSS), guessing that they are better classified as special-use or user-supplied instruments. One (HNO3_GTCIMS) is listed as an ion rate; another (XHNO3) as vdc; a third (HNO3_AERO) as "Aerodyne Nitric Acid".

16. IWC1DC is apparently present only in DC3 and WCR-TEST. It is based on assumptions about ice density, not specified, so it seems best to omit it and classify it "sp" for now. If I can learn what algorithm is used, I can add this.

17. LATF_A/G and LONF_A/G appear to essentially duplicate LAT_A and LON_A, so I have not added what distinguishes them to the descriptions. If I learn more, I should update them.

18. LI8202 and LI8203 appear only in IDEASG and are "Licor Methane Mixing Ratio" so I have typed them and 'sp' and omitted them from PA.

19. MRH2O_CU appears only in IDEAS-4-C130, so it appears to be a special-test instrument and has been classified 'sp' and omitted from PA.

20. I omitted NH3_AERO as likely user-supplied? It doesn't appear in our instrument inventory, and was used only in FRAPPE.

21. NOY_ACD appears occasionally (e.g., WINTER) and NOY_CL appears in START08. Neither lists

any dependencies. The current description of the instrument calls is an no-no2 or no-noy instrument, so it appears that NOY_ACD should be added to the list. WINTER lists NOY (total reactive nitrogen) as a measurement, but all other uses list NO2 as the other product (with NO) from this instrument. I have left NOY_ACD in the appendix, linked to the NO-NO2 description discussed in the following item, and I added NOY_ACD to that description. This may need clarification. I omitted NOY_CL and NOYO3_CL as special-use, and NOY_ARN also; I'm guessing the former are Clemson instruments and the latter was apparently from a NOAA instrument called ARNOLD that was operated by Steve Brown?

22. The final product from the NO-NO2 instrument was listed as XNOCAL/XNYCAL, but I don't find those variables in any netCDF files, production or /scr/raf_data. I did find NO_ACD, NO2_ACD, and FO3_ACD, but the latter is apparently from a different instrument although various instrument descriptions list O3 as another product from the NO-NO2 instrument. I have tentatively changed the description from XNOCAL/XNYCAL to NO_ACD and NO2_ACD. I also changed the units from ppbv to pptv because that is what the instrument description says and those are the units associated with NO_ACD and NO2_ACD when the variables are present in netCDF files.

23. I omitted some variables named O3MR_ with various suffixes (CSD, F1, F2, GMD, CL) because they appeared to be user-supplied. I left their type as 'sp'.

24. PSTRB_06 (260X Total Strobes) apparently hasn't been used since PASE, so I omitted it and left its type is 'st'.

25. REJAT is rejected particle count for the CDP or SPP-100. There is a comment that many status variables like this are omitted, so I have omitted this also.

26. I omitted AQRATIO although it often appears in netCDF files. It is called "Al's Fudge Facter" – not mine, must be a remnant from something Schanot did? It is defined differently in different projects, sometimes depending on ADIFR and QCF, sometimes depending on QCF and PSF; the former seem to match the name (AQRATIO) better, but the latter have continued to be characteristic of the variable up to the present. It apparently was used in an old version of PCORs, last used in FRAPPE in 2014; later projects have no dependencies on AQRATIO. I relegated it to an obsolete variable. It should be omitted from processing where it is not used (and the name is not appropriate).

27. The current description of the humidity-corrected variables for ATX and TASX refer to TASxD and ATxD, but no such variables appear in the netCDF files. Instead, TASDRY is present, and TASFLG=0 denotes when the humidity-corrected values are normal. When TASFLG=1, something is wrong with EWX (usually exceeding the equilibrium value for ATX) so the variable reflects that EWX is replaced by the equilibrium value. It would still be useful to include ATxD using dry-air specific heats and gas constant, but I have omitted this for now because it is not being done. The value for TASDRY is that this would avoid possible discrete jumps at TAFFLG transitions that might otherwise contaminate variance spectra.

28. Regarding humidity-corrected TASX, there should be some dependence on humidity, probably EWX. That should be an additional dependency attributed to MACHX/MACHF.

29. I omitted TKE and classified its type as "sp". I think this is a useful variable and I have argued for its addition, but to date it only appeared in TREX from 2006 so it's not appropriate to include it.

30. The former descriptions listed USFLWC as the corrected UHSAS sample flow, but that variable does not appear in any of the processed archives. Instead, the corrected flow is UFLWC. The variable name has therefore been changed. The units are given in the attributes as sccm (standard cubic cm per minute) but are given in PA as cm^3s^{-1}; I have changed this to cm^3min^{-1}.

31. In TI3GER the dependencies for UFLWC include UVFLIN, not USMPFLW; UVFLIN has no dependencies but apparently refers to a different method of flow control using an Alicat mass-flow controller? The name given for UVFLIN is "Sample (inlet) volume flow calculated from dPinlet" so that would be volumetric flow like USMPFLW, making it reasonable to substitute UVFLIN for USMPFLW in the algorithm? However, units of UVFLIN are "ccs" (should that be sccs?) and typical values are near 1, USMPFLW (sccm) is near 50, so some change in the algorithm is needed. There are no dependencies for UVFLIN and no dPinlet variable, but cal coefficients are given so this must be the direct measurement of flow? The Alicat (a mass-flow meter) probably provides an output representing standard-conditions flow? This needs further clarification and revision. Also, USMPFLW and PFLW have units of "sccm" but UFLWC and PFLWC are in ccs so I have introduced a factor of 60 for this conversion.

32. CUHSAS attributes include a dependence on USCAT but this is not included in the description of the algorithm. This dependence needs to be added to the documentation.

33. VISHB only appears in C-130TEST, with no long-name defined and units of "V". I omitted it.

34. DC3 and NOMADDS include a variable XCO_AL with units ppmv and long_name "Carbon Monoxide". It does not include the expected processed variable from CORAW_AL, COMR_AL, for which units are expected to be ppbv. I omitted XCO_AL for that reason. I didn't find XCO_AL in any other project since 2013. XCOMR has also been used for CO mixing ratio, but in that case it has the expected units of ppbv so I left a reference to it.

35. I omitted XGREFF ("Gerber PV-100 Effective Droplet Radius") as recorded on TEST-C130 because it had units of "V". I kept a reference to XGRFF, which had units of "um", because it also occurred in VOCALS archives.

36. I spent a lot of time trying to get footnotes (that appear at the end of each section) to have superscript numbers, but the bookdown CSS specifies a normal number followed by a period. I learned that this is also the recommendation of the Chicago Manual of Style, so I finally gave up. There is probably a way to do this by modifying CSS.

37. Regarding the VCSEL: CONCV_VXL indicates dependencies on PSXC, ATX, SMPLSRINT_VXL, MODE_VXL, and PREGAIN_VXL. This is unclear to me, and the use of these variables in the processing algorithm needs clarification. The basic measurement is RAWCONC_VXL, which is the sensed molecular density. This needs adjustment to ambient conditions, so the adjustment should at least reference the pressure and temperature in the sensed volume in addition to the ambient pressure and temperature, but those measurements are not included in the dependencies. How are the last three dependencies used. An updated processing algorithm is needed to clarify this.

38. AT_VXL in netCDF files has units of "degK". The proper unit is "kelvin" (with lower-case k) or, abbreviated, "K".

39.