UM | Centre for Earth Observation Science

CanWIN AVOS Cookbook

CENTRE FOR EARTH OBSERVATION SCIENCE

Document Control

0.1 Version History

Version	Author(s)	Type	Date Modified	Comments
1.0	Kerr, L., Iyakoregha, V., Clair, C., Kum, R.	Obsolete copy	2020/06/29	Errors in workflow and cleaning scripts.
2.0	Friesen, K. L.	Working Copy	2022/01/05	Corrected workflow.

0.2 Document Location

A digital copy of this document can be found on GitLab and in the CanWIN datahub.

0.3 License

Friesen, K. L.(2022). AVOS Cookbook, Version 2.0. Centre for Earth Observation Science. University of Manitoba.

Contents

1 AVOS Data

1.1 Critical Summary

This is describes the **data management workflow 1** process for converting AVOS from its raw format.

GPS and meteorological data are collected by the **Automatic Voluntary Observing Ship System (AVOS)** aboard the M/V NAMAO and is saved in a format unable to be easily analyzed by technicians.

1.1.1 Data Management Tools

Converting AVOS data is one step out of a three-step workflow (Figure 1.1) in combining incubator data with GPS and meteorological data; however, this step in the workflow can also be standalone to utilizing AVOS files for other purposes.

1. **AVOS_parse_v1.2 or v1.3:** Converts AVOS .AVMTD files to .csv, converts GPS coordinates to decimal degrees, and subsets raw data to remove unnecessary columns. If running version 1.3 the script standardizes the variables in the converted file.

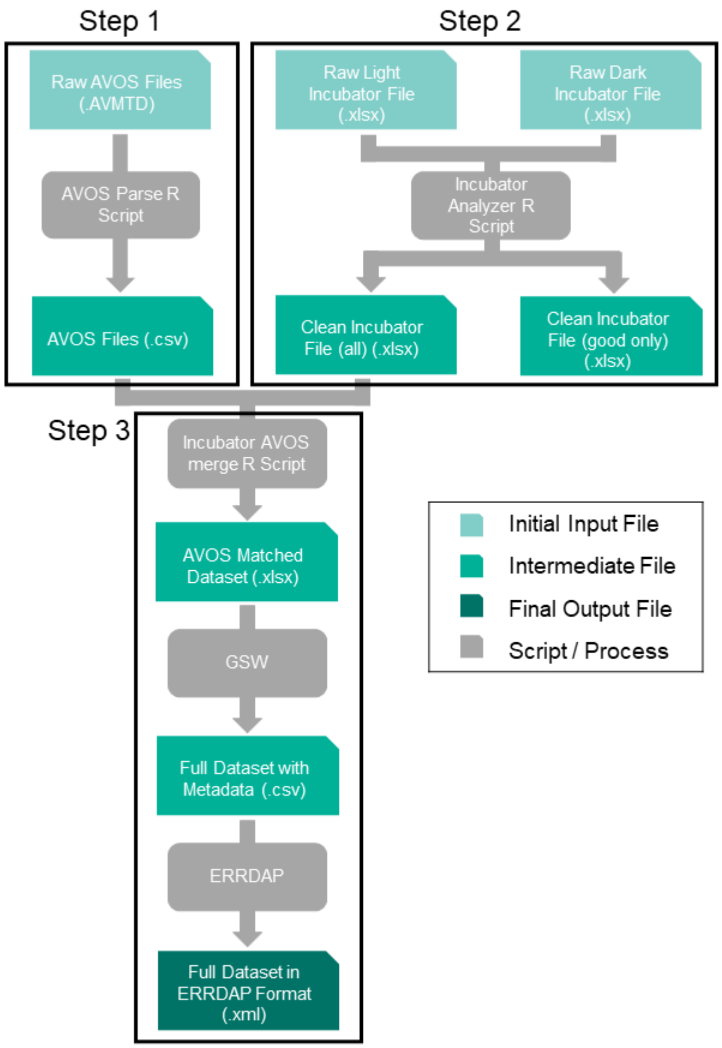

Figure 1.1: AVOS data management step in overall Incubator-AVOS workflow.

1.2 Initial Data

1.2.1 Raw AVOS Files

Dataset Description

GPS and meteorological data is collected continuously by the Automatic Voluntary Observing Ships Systems (AVOS), which is a meteorological station aboard the Namao. AVOS data is recorded at 10 minute intervals, at discrete sample stations, and at multiple depths. Data is received as a collection of **Data Level 0**, comma-delimited .AVMTD files directly from the AVOS computer system.

File Name(s)

AVOS file names: YYYYMMDD.AVMTD

File Source and Location

Raw AVOS files require a request from the **Lake Winnipeg Research Consortium (LWRC)** and there is no current workflow for uploading raw files. Currently, raw AVOS files are available for 2018 and 2019.

Raw AVOS files are stored on GitLab as a 7-zip (.7z) compresssed folder, which can be extracted using the 7-zip software. The folder structure has various .AVMTD files saved per year and per month, as shown below.

- 2017
  - Jan
  - ...
  - Dec
- 2018
  - Jan
  - ...
  - Dec
- ...

Dataset Variables

Table 1.1: Variables in raw AVOS data files

Header	Description	Data Type	Range or Expected Values	Units
V1	Message designator	text	"AVMTD" or "D"	N/A
V2	Sample end date	date	N/A	YYMMDD
V3	Sample end time	time	[000000, 235959]	HHMMSS
V5	GPS latitude	numerical	[-90, 90]	DDmm.mmmmn (n='N' or 'S')
V6	GPS longitude	numerical	[-180, 180]	DDDmm.mmmmn (n = 'E' or 'W')
V7	Apparent wind speed	numerical	[0, 60]	Knots (kn)
V8	Apparent wind direction	numerical	[0, 360]	Degrees (true)
V9	True wind direction	numerical	[0, 360]	Degrees (true)
V13	Barometric pressure, uncorrected	numerical	[500, 1100]	mbar
V15	Air temperature	numerical	[-40, 60]	°C
V16	Relative humidity	proportion	[0, 100]	%
V19	Water temperature	numerical	[-10,50]	°C
V20	Battery voltage	numerical	[12, 110, 220]	Volts Direct Current (VDC)
V23	Ship's call sign	alphanumerical	N/A	N/A
V24	Ship's heading (AVOS magnetic)	numerical	[0, 360]	Degrees
V25	Ship's speed over ground (10-minute average)	numerical	[0, 60]	Knots (kn)
V33	Barometric pressure, corrected to mean sea level	numerical	[500, 1100]	Mbar
V39	Ship's heading (Gyro compass)	numerical	[0, 360]	Degrees

Note: Row names V4, V10-12, V14, V17-18, V21-22, V26-32, V34-38, and V40 are unknown variables and are not relevant to this analysis.

1.3 Scripts and Analytical Processes

1.3.1 AVOS File Conversion

Script Type: R Script.

File(s) In: Multiple raw AVOS files of YYMMDD.AVMTD format.

File(s) Out: Multiple converted AVOS files of YYMMDD.csv format.

User Instructions

To run this R Project, you must have both **R** and **RStudio** installed on your computer.

1. Download **AVOS_parse_v1.2.R** or **AVOS_parse_v1.3.R** from the scripts folder on CanWIN GitLab.
2. Download the raw AVOS (.AVMTD) file(s) from the data folder on the CanWIN GitLab.
3. You will need to extract the files from the zip folder using 7-zip.
4. Open the **AVOS_parse_v1.2.R** script in RStudio.
5. It is recommended that you create the ouput folders for the converted files (detailed in codebook).
6. The output folders should be labelled similarly as the raw AVOS folders/directories, except when listing the months you should use the numerical representations (e.g. 2018/12 instead of 2018/Dec).
7. Place cursor at the indicated location in the script and click "Run" at the right-hand side of the code window in RStudio. This process should be completed within a minute.
8. Your output files should now be available in the directory you specified in the dialog box, as well as the path printed in the Console window of RStudio as a reminder.

Back-End Details

Back-End Script Location: There are no back-end scripts running for this process.

Libraries Used: R package used in script are svDialogs.

Analytical Process

1. After clicking "Run", a dialog box appear on your screen requesting the input directory of the raw AVOS files.
2. If no pathway is provided the script will stop.

3. The script will also stop if the directory you pointed the script to doesn't contain the .AVMTD files.

4. Another dialog box appears to request the output directory where converted AVOS .csv files are to be saved.

5. If no pathway is provided the script will stop.

6. The script retrieves and loops through the raw AVOS files and prints the names of the files as they are processed.

7. Converts datetime to readable format and saves year, month, date, hour and minutes in new columns.

8. Converts latitude and longitude to decimal degrees.

9. Retrieves remaining variables from raw AVOS file.

10. Merges date, time, latitude, and longitude with remaining variables in new dataset.

11. Saves new dataset to the specified output directory as a .csv.

12. At the end of the conversion script, a message will print in the console indicating the end of the process and giving the directory path where the .csv files are saved.

NOTE: Since you can only point the script to a month of .AVMTD files you only will ever have the capacity to convert AVOS files for a given month. This process will need to be repeated if you require more than one month out of a year. Hence, creating output directories with the year (YYYY) and month (MM) to organize your converted files (e.g. 2018/12/20181204.csv) will be helpful in continuing on to the next steps of the workflow (Figure 1.1) script which assumes and searches for this folder structure when merging incubator and AVOS files.

1.4 Final Output Files

1.4.1 Converted AVOS Files

Dataset Description

The collection of converted AVOS files differ only by day and contain measurements for every hour of the day at intervals of 10 minutes. Please ensure you create the file structure as recommended in the previous instructions if you are progressing to step 3 (Figure 1.1) for the final merge process.

File Name(s)

AVOS file names: YYYYMMDD.csv

File Source and Location

These files will be saved to your local drive to the pathway you specified when running the **AVOS_parse_v1.2.R** script. A print message will indicate the end of processing and give the output directory path where the .csv files are located.

Dataset Variables

Converted files are NOT automatically sorted into nested folders by year (YYYY) and month (MM). You will need to create the structure to reflect the raw folder structure, which lists months by name instead of numerical value. The output pathway structure should be structured as follows:

- 2017
  - 01
  - ...
  - 12
- 2018
  - 01
  - ...
  - 12
- ...

Individual files contain the following variables:

Table 1.2: Variables in converted AVOS data files (v1.2)

Header	Description	Data Type	Range or Expected Values	Units
year	Sample year	numerical	[2017, 2019]	N/A
month	Sample month	numerical	[01, 12]	N/A
day	Sample end day	numerical	[1, 31]	N/A
hour	Sample end hour	numerical	[0, 23]	N/A
minute	Sample end minute(s)	numerical	[0, 50]	N/A
lat	GPS latitude	numerical	[50.0000, 50.9999]	Decimal degrees (dd)
lon	GPS longitude	numerical	[-96.0000, -96.9999]	Decimal degrees (dd)
ws	Apparent wind speed	numerical	[0, 60]	Knots (kn)
wd_rel	Apparent wind direction	numerical	[0, 360]	Degrees (true)
wd_true	True wind direction	numerical	[0, 360]	Degrees (true)
P_uncor	Barometric pressure, uncorrected	numerical	[500, 1100]	mbar
Ta	Air temperature	numerical	[-40, 60]	°C
RH	Relative humidity	proportion	[0, 100]	%
Tsrfc	Water surface temperature	numerical	[-10,50]	°C
sog_kts	Ship's speed over ground (10-minute average)	numerical	[0, 60]	Knots (kn)
P_cor	Barometric pressure, corrected to mean sea level	numerical	[500, 1100]	Mbar
heading	Ship's heading (Gyro compass)	numerical	[0, 360]	Degrees

For files going through the standardization process in version 1.3, individual files contain the following variables:

Table 1.3: Variables in converted AVOS data files (v1.3)

Header	**Description**	**Data Type**	**Range or Expected Values**	**Units**
Date_and_Time	datetime	numerical	NA	YYYY-MM-DD hh:mm
latitude	GPS latitude	numerical	[50.0000, 50.9999]	Decimal degrees (dd)
longitude	GPS longitude	numerical	[-96.0000, -96.9999]	Decimal degrees (dd)
wind_speed	Apparent wind speed	numerical	[0, 60]	Knots (kn)
RelWindDirFrom	Apparent wind direction	numerical	[0, 360]	Degrees (true)
WindDirFrom	True wind direction	numerical	[0, 360]	Degrees (true)
air_pressure	Barometric pressure, uncorrected	numerical	[500, 1100]	mbar
air_temperature	Air temperature	numerical	[-40, 60]	°C
relative_humidity	Relative humidity	proportion	[0, 100]	%
Temp	Water surface temperature	numerical	[-10,50]	°C
platform_speed_wrt_ground	Ship's speed over ground (10-minute average)	numerical	[0, 60]	Knots (kn)
air_pressure_at_mean_sea_level	Barometric pressure, corrected to mean sea level	numerical	[500, 1100]	Mbar
platform_orientation	Ship's heading (Gyro compass)	numerical	[0, 360]	Degrees

A Reference Tables

A.1 Data Levels

Level 0 – Raw data: unprocessed data and data products that have not undergone quality control. Depending on the data type and data transmission system, raw data may be available within seconds or minutes after real-time. Examples include real-time precipitation, streamflow, and water quality measurements

Level 0.1 – First pass QC: A first quality control pass has been performed to remove out of range and obviously erroneous values. These values are deleted from the record. E.g: Online Environment Canada stream-flow data, laboratory data

Level 1 – Quality Controlled Data: Data that have passed quality assurance procedures such as Level 0.1 and have been further quality controlled by data provider before being submitted to CanWIN (e.g. Idronaut data with only downwelling (upwelling data removed) data included.

Level 1.5 – Advanced Quality Controlled Data: Data have undergone complete data provenance (i.e. standardized) in CanWIN. Metadata includes links to protocols and methods, sample collection details, incorporates CanWIN's or another standardized vocabulary, and has analytical units standardized. Note: Process still under development in CanWIN (as of May 13, 2020).

Level 2 – Derived Products: Derived products require scientific and technical interpretation and can include multiple data types. E.g.: watershed average stream runoff derived from stream-flow gauges using an interpolation procedure.

Level 3 – Interpreted Products: These products require researcher (PI) driven analysis and interpretation and/or model-based interpretation using other data and/or strong prior assumptions. E.g.: watershed average stream runoff and flow using streamflow gauges and radarsat imagery

Level 4 – Knowledge Products: These products require researcher (PI) driven scientific interpretation and multidisciplinary data integration and include model-based interpretation using other data and/or strong prior assumptions. E.g.: watershed average nutrient runoff concentrations derived from the combination of stream-flow gauges and nutrient values.

Content retrieved from https://lwbin.cc.umanitoba.ca on July 06, 2020.

A.2 Result Value Qualifiers

ADL	Above Detection Limit
BDL	Below Detection Limit
FD	Field Duplicate
LD	Lab Duplicate
$	Incorrect sample container
EFAI	Equipment failure, sample lost
FEF	Field equipment failed
FEQ	Field Equipment Questionable
FFB	Failed. Field blank not acceptable
FFD	Failed. Field Duplicate
FFS	Failed. Field spike not acceptable
H	Holding time exceeded
ISP	Improper sample preservation
ITNA	Incubation time not attained
ITNM	Incubation temperature not maintained
JCW	Sample container damaged, sample lost
NaN	Value is missing and reason is not known
NC	Not collected
ND	Not detected
NR	Sample taken/measured on site but information in this field not recorded
NS	Sample collected but not submitted
OC	Master Coordinate List Used
P	Analysis requested and result pending

prob_good probably good value. Data value that is probably consistent with real phenomena but this is unconfirmed or data value forming part of a malfunction that is considered too small to affect the overall quality of the data object of which it is a part

prob_bad probably bad value. Data value recognised as unusual during quality control that forms part of a feature that is probably inconsistent with real phenomena

Interpolated This value has been derived by interpolation from other values in the data object

Q Below limit of quantification (LOQ). The value was below the LOQ of the analytical method. The value in the result field is the limit of quantification (limit of detection) for the method

B Glossary of Options and Packages

B.1 R Packages

Visit https://cran.r-project.org/web/packages/available_packages_by_name.html to learn more about R packages

- **Package 1** - Description
- **Package 2** - Description

B.2 Python

B.2.1 Python Script-Specific Options

- **Option 1** - Description
- **Option 2** - Description

B.2.2 Python Packages

Visit https://docs.python.org/3/library/ to learn more about python packages

- **Package 1** - Description
- **Package 2** - Description

Example: Section 2.1 from Victory's semi-hemi codebook