Stable isotope analysis (SIA) is a powerful tool that has been a staple of the fields of oceanography, geochemistry, and forensic sciences. Recently, the application of SIA has expanded to other fields such as ecology and animal biology. However, the incorporation of SIA to ecology and animal biology is often limited to carbon and nitrogen isotopes. In addition, many researchers in the fields of ecology and animal biology choose to send their samples to other institutions for analysis instead of completing their analyses in-house because of the expenses and logistics related to traditional isotope analysis. The development of new instrumentation such as cavity ring-down spectroscopy (CRDS) and off-axis integrated cavity output spectroscopy provide a potential solution to these setbacks. In particular, CRDS is an affordable alternative to the long-standing traditional reliance on isotope ratio mass spectrometry (IRMS). CRDS instruments, such as the Picarro L2140-*i*, provide high-precision measurements of hydrogen and oxygen stable isotopes in a fraction of the time required for IRMS analysis, and CRDS instruments possess automated functionality, allowing for a less labor-intensive process compared to IRMS.

Stable isotopes of oxygen and hydrogen (measured as δ17O, δ18O, and δ2H) can provide information related to the sources of environmental water intake and to the animal’s body water pool. Most hydrogen and oxygen analyses use distilled blood plasma or serum samples, but under the right circumstances (e.g., a trained animal in captivity) distilled saliva and urine samples can be collected as well. A central premise for all these methods is that water is critical to animal biology. Most terrestrial animals are ~60-70% water by mass, and this body water comes from a combination of environmental sources and endogenous processes (e.g., newly-synthesized metabolic water, a byproduct of metabolic pathways.

The code stored in this repository (file name “Capstone\_Project”) is designed to help expand the use of stable isotope analysis via CRDS to the fields of ecology and animal biology. While CRDS instruments such as the Picarro L2140-*i* are user-friendly water isotope analyzers, the concepts related to SIA are still daunting and may be intimidating for many ecologists or animal biologists who may consider incorporating this type of instrumentation into their lab. Many researchers continue to rely on shipping their water or plasma samples to outside institutions for analysis due to this obstacle. The purpose of this repository is to provide simple and efficient code to demonstrate the ease with which a researcher can correct and finalize data from a CRDS water analyzer instrument for publication. For example, while the raw datasheet exported from a CRDS instrument like the Picarro L2140-*i* provides data related to >25 variables, the variables of interest are as follows:

1. δ17O (in the code as ‘d(17\_16)Mean’ for raw values and ‘d17O\_amended’ for corrected)
2. δ18O (in the code as ‘d(18\_16)Mean’ for raw values and ‘d18O\_amended’ for corrected)
3. Δ17O (calculated from δ17O and δ18O; in the code as ‘E17\_Mean’ for raw values and ‘E17O\_amended’ for corrected)
4. δ2H [deuterium;2H] (in the code as ‘d(D\_H)Mean’ for raw values)
5. average water concentration (in the code as ‘H2O\_Mean)

Ultimately, the code provided in this repository generates an exportable datasheet with just these variables of interest.

This repository contains six datasets generated from my lab: 3 datasets related to ‘standards runs’ and 3 datasets related to ‘analysis runs’. Each standards run dataset is composed of ~600 measurements, while analysis run is composed of 300-450 measurements. The standards run datasets involve measurements of internationally accepted water standards (VSMOW [Vienna Standard Mean Ocean Water] and SLAP [Standard Light Antarctic Precipitation Water], along with other in-house waters. The standards run is designed as a template so that isotope values of the waters corrected by measurements of VSMOW and SLAP during this run can be established for the next set of analysis runs which also contain samples from different animal plasma with unknown values. For example, a completed 3-day standards run will influence the next 2-months of analysis runs. At the end of these two months, another standards run will be conducted to establish new values for in-house waters and the cycle repeats. Once a standards run is completed, the raw data should be exported and then imported into python using the provided. The code then builds a correction equation by comparing the known and measured values of the VSMOW and SLAP for δ17O and δ18O. The individual correction equations are then applied to the raw δ17O and δ18O values of each water included in the standards run, allowing for generation of corrected values for δ17O, δ18O, and Δ17O. The corrected values are then compared to known and established in-house values to ensure that the run was successful. The code provided should effectively cycle through a standards run, produce comparable isotope values to the known and in-house established values, and then these corrected, measured values can be used to successfully inform proceeding analysis runs that use in-house standards (such as USGS47 and USGS50 [United States Geological Survey]) to correct values instead of VSMOW and SLAP (since these waters are expensive and only available in small volumes; see an example of an analysis run via Table 1).

To provide support for this, I have provided three raw datasheets for both standards runs and analysis runs. Each standards run datasheet is paired with an analysis run datasheet (e.g., ‘StandardsRun\_1’ is compatible with ‘AnalysisRun\_1’). Each pairing should demonstrate the following:

1. That generating a correction equation from VSMOW and SLAP measurements during standards runs should generate:
   1. δ17O values within 0.15‰ (per mi) of the known/established value
   2. δ18O values within 0.3‰ of the known/established value
   3. Δ17O values within 0.015‰ of the known/established value
2. That the finalized values from the standards run should be incorporated into the data environment such that it can be used to influence the proceeding analysis run
3. That generating a correction equation from in-house water standards (in the provided code these are ‘USGS47’ and ‘USGS50’) with values established from the preceding standards run should generate:
   1. δ17O values within 0.15‰ of the known/established value
   2. δ18O values within 0.3‰ of the known/established value
   3. Δ17O values within 0.015‰ of the known/established value

The final datasheet that is provided from the code for both the standards run and analysis run includes a comparison section that allows for easy determination if the following data benchmarks are met.

This code is designed to be as fully automated as possible. As such, only the following sections should require editing: 1) Section 3 for importing raw Standards datasheet; 2) Section 37 for determining where to export csv files to; 3) Section 43 to set which conditioning water samples to remove; 4) Section 46 to determine where to export xlsx file to; 5) Section 48 for importing raw Standards datasheet; 6) Sections 63-70 depending on what in-house water standards are used by the lab; 7) Sections 78-79 for determining where to export csv files to and to remove the two in-house standards from the list; 8) Section 83 to set which conditioning water samples to remove; and 9) Section 87 to determine where to export xlsx file to.

FILES

Capstone\_Project. .ipynb – Jupyter notebook file containing the Python code for correcting and finalizing standards and analysis run raw data

StandardsRun\_1.csv – CSV file containing raw data from a standards run completed on 30-January-2023. This file should be imported into the code via Section 3 of the code. This csv should only be used in combination with the AnlaysisRun\_1 file.

StandardsRun\_2.csv – CSV file containing raw data from a standards run completed on 26-Feb-2023. This file should be imported into the code via Section 3 of the code. This csv should only be used in combination with the AnlaysisRun\_2 file.

StandardsRun\_3.csv – CSV file containing raw data from a standards run completed on 25-Sept-2022. This file should be imported into the code via Section 3 of the code. This csv should only be used in combination with the AnlaysisRun\_3 file.

AnalysisRun\_1.csv – CSV file containing raw data from an analysis run completed on 02-Feb-2023. This file should be imported into the code via Section 48 of the code. This csv should only be used in combination with the StandardsRun\_1 file.

AnalysisRun\_2.csv – CSV file containing raw data from an analysis run completed on 01-Mar-2023. This file should be imported into the code via Section 48 of the code. This csv should only be used in combination with the StandardsRun\_2 file.

AnalysisRun\_3.csv – CSV file containing raw data from an analysis run completed on 29-Sept-2022. This file should be imported into the code via Section 48 of the code. This csv should only be used in combination with the StandardsRun\_3 file.

Requirements.txt – Text file containing the python and package requirements to run the code.

OUTPUTS

csv.files – For each standards and analysis run, csv files will be generated for each individual water sample analyzed during the run. These csv files contain all the variables measured by the CRDS instrument and are mainly generated as a backup source if further analysis was ever required. These files are generated in the following format: SampleName\_Day\_Year\_Month.csv

.xlsx files – For each standards and analysis run, a single xlsx file will be generated that contains all the finalized data for each water sample as a sheet within the xlsx file. These sheets only contain the variables of interest and are meant to be used for publication and for generating larger compiled datasheets. These files are generated in the following format: Standards\_Day\_Year\_Month.xlsx or Analysis\_Run\_Day\_Year\_Month.xlsx.

TYPES OF WATERS FOR ANALYSIS

Internationally accepted water standards – VSMSOW (Vienna Mean Standard Ocean Water) and SLAP (Standard Light Antarctic Precipitation Water)

In-house water standards – USGS47 and USGS50 (Waters provided by the United States Geological Survey for interlaboratory comparison)

Control waters – USGS46, USGS48, Real\_KD (Kona Deep brand water), SeatW (Tap water from Seattle, WA), MissMT (Tap water from Missoula, MT), IceLava (Iceland Lava brand water), SupremeBoil (water boiled for 3 days to enrich δ18O value), MegaBoil (water boiled for 5 days to enrich δ18O value)

Conditioning vials – DummyKD, SuperBoil, UltraBoil, USGS46.25, USGS46.5, USGS46.75, SLAP.25, SLAP.5, SLAP.75 (conditioning vials should be removed via the Python code during the process)

**Table 1**.Example of an analysis run completed on the Picarro L2140-*i*. Two in-house standards (VA01 and VA02) verified via internationally-accepted water standards were incorporated throughout the run to enable correction of raw measurements. Injection total refers to the total number of injections as the run progresses while number of injections refers to the number of injections for that particular item, with more injections required for analyzed samples that vary greatly in δ18O from the proceeding item. Autosampler job number refers to the order in which each item was listed for analysis. While a sample may have 30+ measurements, most of these measurements are removed due to the memory effect and only a select number of measurements remain for analysis.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Material being analyzed | Injection total | δ18O (‰) | Number of injections | Autosampler job number | Purpose |
| Conditioning vial | 1-70 | ~ 0 | 70 | 1 | Warm-up instrument |
| Conditioning vial | 71-97 | ~ +4 | 27 | 2 | Positioningδ18O for in-house standard analysis |
| In-house standard (VA01) | 98-157 | ~ +8 | 60 | 3 | First of three points for correction equation |
| Control vial | 158-184 | ~ +4 | 27 | 4 | Further validation of data |
| Control vial | 185-211 | ~ -2 | 27 | 5 | Further validation of data & positioning δ18O for unknown samples |
| Distilled  Samples | 212-361 | ??? | 150 | 6-17 | Assessing isotopic values of unknown biological samples |
| Conditioning vial | 362-394 | ~ -4 | 33 | 18 | Positioning δ18O after running unknown samples |
| Control vial | 395-437 | ~ -10 | 43 | 19 | Further validation of data & positioning δ18O for in-house standard |
| In-house standard (VA02) | 438-470 | ~ -9 | 33 | 20 | Third point for correction equation |
| *Total Injections* | *470* |  |  |  |  |
| *Estimated run time* | *~ 72 hours* | |  |  |  |