This is documentation provided by creator PI Rachel Stanley, edited to indicate the ultimate data entities to be provided in EDI packages.

Rachel’s goal: “I need raw data for TOI and O2/Ar. And I need calculated rates for discrete and continuous. So 4 samples to give them.”

Package 6: Raw EIMS TOI concatenated cruises

Described data table csv files:

* High frequency: nes-lter-eims-transect.csv (raw EIMS O2/Ar underway continuous)
* Low frequency: nes-lter-toi-transect.csv (raw TOI rosette and underway discrete)

Corresponding other entities:

* High frequency: **RaEn617withbiosat.mat (3)**
* Low frequency: **bottleEn617withoutincubation.mat (1)**

Package 7: Rates NCP GOP cruise EN617

Described data table csv files:

* High frequency: nes-lter-ncp-transect-summer-2018.csv (rate NCP derived from raw EIMS)
* Low frequency: nes-lter-gop-transect-summer-2018.csv (rate GOP derived from subset raw TOI only at the surface)

Corresponding other entities:

* High frequency: **ncplterEn617.mat (4)**
* Low frequency: **discreteratesEn617.mat (2)**

1. Raw data for TOI:

Even the raw data can be matlab files since then I can include the temperature and salinity if they want it.

As is, it just has the following:

bottleEn617withoutincubation.mat

Variable is “d”

columns are 1: date in matlab format 2: delta O2/ar 3: Ratio O2/ar 4: depth 5. D17 6. d17 7. d18 8. Niskin number. NOTE – underway samples are listed as depth of 0 and Niskin of 0.

depthcutoff=6; % depth in m which we consider "surface"

1. GOP rates: (only at surface – from underway or surface CTD)

File is discreteratesEn617.

Variable is A.

Columns are 1. time, 2. lat, 3. lon, 4. gop, 5. ncp, and 6. ncp/gop

Units of GOP and NCP are mmol O2/m2/day and reflect the productivity integrated over the mixed layer

1. Data for O2/Ar

This data has already had some initial calibrations applied – I have calibrated the data according to the ratio measured in air roughly every 6 hours and have removed the data when the mass spec was measuring air – i.e. it gives a calibrated O2/Ar ratio that can then be used to calculate rates of NCP. This data also includes the biological saturation – the O2/Ar divided by the equilibrium value of O2/Ar – this gives the expected super (or under) saturation of O2 based on biological factors anyway. A calculation that went into it is the solubility of T and S that depends on the T and S data so in some sense, maybe it should be in derived data since that number could change if T or S data changes. But it is straightforward to calculate (with many fewer issues than NCP) and is much more meaningful to people than just O2/Ar so I think it would be nice to include it. We can include the T and S we used so people can recalculate if they want based ont eh O2/Ar raw ratio and whatever T or S they want.

At the moment, this is variable called RaEn617withbiosat.mat

Variable is Ra

Columns are 1. Matlab date. 2. O2/Ar corrected for air values. 3. Water temp. 4. salinity 5. Latitude 6. Longitude. 7. Cumulative distance.

The cumulative distance we probably should not include since it may not be accurate (was a bit of an approximate algorithm I used). But the files I send to you will usually have it since that is what my code spits out. Of course I can remove it for you before sending if you’d like.

1. Data for NCP: ncplterEn617.mat

Variable is ncp. Units of ncp are mmol O2/m2/day and it just reflects NCP integrated over the mixed layer. The columns are

1. matlab date 2. O2/Ar corrected for air values 3. temp 4. salinity.

5. lat 6. lon 7. cum dist 8. biosat 9. ncp 10. k

So really this is just the first 8 columns of the O2/Ar raw data (point 3) plus ncp and k. NCP is the rate of net community production integrated over the mixed layer. K is the gas transfer velocity that was used when calculating ncp from the O2/Ar ratio. K has the biggest uncertainty of any of the terms and the most different ways the community addresses it/calculates it.

If it would be helpful, I can add in a column for O2 saturation which is another term that goes into the calculating of NCP (it is based on T and S so I don’t include it separately as a column but since there are different algorithms used, maybe it is best I do so?)