SFER Data Changes

In 2024, the South Florida Ecosystem Restoration (SFER) cruise dataset underwent an overhaul to clean the data, make it machine-readable, allow for flexibility in extra sampling, and allow for easier hosting of the data. During this overhaul several changes withing the data were made. Here is a list of changes that occurred.

Cruise ID

The cruise ID consists of the ship's initials, the last two digits of the year, and the Julian day that sailing begins. For consistency and cleanliness, all ship initials were changed to two letters. Cruise IDs that did not contain the proper formatting were corrected for consistency.

Table 1: Cruise ID Changes

Old Cruise ID	New Cruise ID	Changes made
WS1418	WS14335	Number ID changed to Julian day
SAV1803	SV18067	Ship initial changed to two letters; number changed to Julian day
SAV18173	SV18173	Ship initial changed to two letters
WS20278	WS20279	Julian day corrected
H23138	HG23138	Ship initial changed to two letters

Ship IDs

The research vessel is represented in the cruise ID with the ship's initials. Below are the abbreviations and the institution the ship is associated with.

Table 2: Ship ID abbreviations

Ship Abbreviation	Research Vessel	Institution
WS	F. G. Walton Smith	University of Miami
WB	R/V Weatherbird II	FIO
HG	R/V Hogarth	FIO
SV	R/V Savannah	Skidaway

Station Names

Some station names were changed to provide consistency within the entire data set (i.e. if a station name had changed over the years and had multiple identifiers), and/or to be machine readable. The notation of the letter "B" or "b" for stations reoccupied on a cruise were also removed; the stations are uniquely identified by the addition of cast number for or date (see cast number section for more details).

Table 3: Station name changes

Old Station Name	**New Station Name**	**Changes Made**
21/LK	21LK	Machine readable; inconsistencies were also corrected (historically some stations labeled as 21 or LK)
54B, 55B, 56B, 57B, 57.1B, 57.2B, 57.3B	54, 55, 56, 57, 57.1, 57.2, 57.3	
8b, 9B	8, 9	
BG7B, BG13B, BG15B, BG17B, BG20B	BG7, BG13, BG15, BG17, BG20	Differentiated by cast number
CAL-A	CALA	Machine readable
EK IN, EK_IN	EKIN	Machine readable
EK MID, EK_MID	EKMID	Machine readable
EK OFF, EK_OFF	EKOFF	Machine readable
Naples Blue hole	NaplesBlueHole	Machine readable
Red Tide	RedTide	Machine readable
TB-A, TB-B, TB-C	TBA, TBB, TBC	Machine readable
UK IN, UK_IN	UKIN	Machine readable
UK MID, UK_MID	UKMID	Machine readable
UK OFF, UK_OFF	UKOFF	Machine readable
Z03-xxx, Z04-xxx	Z03xxx, Z04xxx	Machine readable for full range of numbers

Station Type

Station type is differentiated as a CTD cast (C) or flowthrough sampling while underway (F). Changes were made to make these differentiations consistent within the dataset.

Table 4: Station type changes

Old Station Type Name	New Station Type Name	Changes Made
CTD	C	Some station types were denoted as CTD for some cruises; changed to C to make entire dataset consistent

Sampling Time

Sampling time is recorded as GMT in a 24-hour format. Changes were made to time formatted in 12-hour format (all AM/PM removed).

Cast Number

A column denoting cast number has been added to differentiate consecutive casts at the same station. If a station is occupied more than once during a cruise, the cast number will reset for each occupation. For example, if you stop at a station, cast the CTD, process samples, and cast another CTD, that cast numbers will be 1 and 2. However, if you stop at a station, cast the CTD, leave the station, come back to that station later that day, then both numbers will be 1. Flow through (F) stations will have a value of "0" in the cast number column.

Historically, stations along the Shark River line (54-57.3) were often occupied twice during a cruise, with the second occupation denoted with a "B" appended to the station name. The letter appendage has been removed (see Station Names section); the sampling events are differentiated by date-time. The cast number for stations like this is 1 for each cast.

Old Column Names -> New Column Names

To make the dataset machine_readable, column names were switched from the previous format. This is not the final list of columns in the dataset, and a final dataset metadata file is available here (TBD). However, this list will help you better understand some of the changes from the old data to the new.

Table 5: Old column names to new

Current Column Name	New Column Name (machine readable)
Rank	rank
Keyfield	keyfield
Cruise	cruise_id
Date (GMT)	date
Temperature	temp

Current Column Name	New Column Name (machine readable)
Salinity	sal
Latitude Deg	lat_deg
Latitude Min	lat_min
Latitude Decimal	lat_dec
Longitude Deg	lon_deg
Longitude Min	lon_min
Longitude Decimal	lon_dec
Time (GMT)	time
Station	station
F or CTD	station_type
Depth	depth
Chlorophyll Tube #	chl_tube
Avg chl a (ug/L)	avg_chla
Avg Phaeophytin (ug/L)	avg_phaeo
Nutrients Tube #	nuts_tube
NH4 (uM)	nh4
PO4 (uM)	po4
NO3+NO2 (uM)	no3_no2
NO3 (uM)	no3
NO2 (uM)	no2
Si (uM)	si
DIC Bottle # Manzello/Enochs	dic_bottle_enochs
TA with CRM Correction (uequiv/kg)	ta_enochs
CRM Corrected TCO2 (µmol/kg)	tco2_enochs
DIC Bottle # Wanninkhof	dic_bottle_wanninkhof
DIC (umol/kg)	dic_wanninkhof
pH	ph_wanninkhof
Oxygen (umol/L)	o2
Oxygen (mg/L) from CTD logs	o2_ctd
Notes	notes