dbSEABED Output Format Specifications for Pointwise Data

Chris Jenkins

(chris.jenkins@colorado.edu)

INSTAAR, University of Colorado

Last edited: 21 Apr 2021

General Description

System-Wide Concepts

dbSEABED outputs many types of data products. The basic output however, is a set of comma-delimited text tables which can be imported into almost all GIS, Relational Database, Spreadsheet and Maths applications. Those files are comma delimited ASCII with Intel ('PC') bit arrangement. They are a very widely acceptable “lowest common denominator” form of data.

The output parameters are classified into “positioning”, “housekeeping”, and “attributes”. Positioning gives information on the locations, depths and times of observations. The housekeeping describes the structure, operation and usability of the data for users. The attributes (“seabed parameters”) give the character, description and properties of the seafloor.

dbSEABED is an Information Processing System - not a Relational Database (RDB) or a Geographic Information System (GIS) - though it is able to generate standard RDB and GIS products. It is a Data-Software-File system, not the most sophisticated technology, but achieving very high levels of usability amongst domain scientists and ocean managers. Higher-technology products can be made from the standard deliveries.

The processing is a stream of processes (Fig. 1) with more advanced and specialized products being built further along or at branches in the processing. The primary products are CSV listings of point data, and/or ASC gridded data according to many projections. Branches in the processing allow for specialized products to be computed at various stages. Those specialized outputs - which will not be detailed here - are for use with netCDF, Google Earth, VRML CoreNavigator, and Rockware.

Usual Delivery (2021)

During 2021 the usual delivery to external projects includes these files. The suffix “f” denotes the final, integrated data format.

Check the details of the parameters and formats in the main text below.

File Name	Description
***_FNLf	This is the data for the ‘top-20’ parameters in its final condition, for use in all projects. *This file format is the most highly useable of all.*
***_SRC	Information about the sources of the data, data entry methods, releasability, and its extent
***_CMP / _CMPf	Components and their abundances, actually also features and their intensities of development. It lists the Fuzzy Memberships (as percent) of many important features and components of the seabed, notably the mineralogy and skeletonized biota.
***_FAC / _FACf	Facies classification in terms of components and features.
***_CLRf	Compositional Analysis (CoDA) logratio transformed values for gravel:sand:mud.

Output data formats

Field arrangement

A header identifying the fields (columns) occupies the first line of each table. All Header entries are in quotes (""). All freeform string data are also enclosed in quotes (""), but some codes that are used in databasing are not. The data in each column is the same data-type for all samples.

For numerics -99 is the usual “Null” or 'No Data' value; for strings it is "-". An exception is made for Latitudes and Longitudes; where 'Null' is -999.

It is important to understand why the data are presented the way they are. Tables have been chosen over XML and over distributed normalized database tables because the users of dbSEABED are familiar and comfortable with GIS tabled data and spreadsheets. Most users are domain scientists, not database managers. We want to optimize the usefulness of dbSEABED, and technical finesse comes second.

Linking the tables

As the various tables are output, datasets, sites and samples of the information are assigned database keys. Examples are “144567”, and “au9:144567”. The keys link across the various tbles, for instance the tables of primary properties, components, facies, data sources, run diagnostics, and their derivatives. The number sequence of keys may have gaps, but the keys are unique across the system.

Since dbSEABED is constantly being updated, a different key may be assigned to a particular sample in successive versions of the system.

Nulls

A very important driver of the overall design is the fact that so many fields are Null. The data are extremely sparse. To help manage this the table structure as a whole is partially normalized, so that a sample may be represented in some tables and not appear others, depending on the outputs. If all fields in a table are null the sample is omitted.

Arrangements for nulls and the marking of all string data with quotes, were driven by the need to control irregular behaviors of some databases whenever data in a column was not universally marked the same way (dBase and MS Access particularly). (Some databases sense null formats using a small initial sample of the data.)

Top 20 Attributes

The Primary Table structure presents the essential navigational, operational, and descriptive data in a highly integrated and fairly disciplined way, again with an emphasis on usability. The same “Top-20” of attributes is given for the extracted, parsed, and on-calculated outputs which are put to separate files because they have different reliabilities and users may wish to judge for themselves which to use.

Actually, most users will wish to combine them because using all 3 reduces the biggest uncertainty in seabed mapping – spatial gaps between sample stations. It also produces the best, least sample-biased final result, especially when descriptive (word) and analytical (numerical) data are combined.

Sample Absences / Data Filtering

Some samples in the raw data collection will fail to appear in outputs because of quality issues, the current state of parsing descriptions, or because they have data which is not represented in the “top-20” parameters.

For those samples refer to the raw data. In seabed cores, samples within the core may fail to appear even though their neighboring samples in the core do appear. This is for the same reasons (quality filters). The dropping-out can be awkward, but is necessary. Over time, it is cured as datasets become better controlled in quality. To achieve total representation of inputs in outputs, the original data must be revised, lifted in quality, and then reprocessed, a process that is aided by the run-diagnostics which are also output.

Formats Overview

Column Groupings

The output parameters span these themes:

GROUP	TYPE	EXPLANATION
Positioning	Coordinates	Location, Water depth, Subbottom depths, Date/Time
Housekeeping	Data Organization	Database keys, Description of sampling methods, Audit codes, Details of selective- or sub-samplings
Attributes	Texture	Sediment fractions, grainsizes and sorting
	Classification	Coded assignment to a sediment facies or habitat class, geologic age, colour
	Composition	The chemistry or mineralogy
	PhysicalProperties	Geotechnical, geoacoustic or physical state

The Positioning and Housekeeping fields are repeated at the front of almost all the files (exceptions: SRC, DGN).

Relational Database Keys

All records in all the files are relationally keyed using sequential numeric indexes for (i) the dataset, (ii) the site, and (iii) the sample/observation. These indexes will change with each new run; they are set for each data delivery to users. This has to be the case because of edits and improvements to the input datasets, processing software, and dictionaries between each delivery.

Filename Formats

Output filenames are usually named "***_XXX.txt" where *** is a prefix set by the operator to denote a data region or source, for example "au9" for the Australian region, "pan" for PANGAEA-sourced data. “XXX” denotes the type of table and tells the table’s derivation and function. Some files are ”8.3”-limit formatted names to work with the QBasic64 legacy software.

(Pre-2011, filenames tend to be put as upper case, to prepare for the fact that some FTP applications change filenames to upper on transfer to UNIX web sites, where filename case is important.)

The “f“ suffix such as in “***_PRSf.txt” signifies that finalization of the format ready for full use in GIS, etc., with: infilling of missing water depths, making system-wide keys.

Catalogue of Output Files

The table types and names are as follows.

XXX	DETAILS
	STAGE 1 OUTPUTS (Initial processing**)**
***_EXT / EXTf	Extracted Data (no significant alteration of input data), usually numeric or coded data and usually based on an instrumental analyses (probe or laboratory) or calibrated observation.
***_PRS / _PRSf	Parsed data, from word-based descripive inputs. This data is less precise than EXT, but includes indispensible extra information on outsized elements, the biota, structures, odours, consolidations, etc.
***_CLC / _CLCf	On-Calculated, Estimated, Modelled data based on EXT or PRS outputs. Estabilished theoretical or empirical functions are used.
***_SRC	Information about the sources of the data, data entry methods, releasability, and its extent
***_CMP / _CMPf	Components and their abundances, actually also features and their intensities of development. It lists the Fuzzy Memberships (as percent) of many important features and components of the seabed, notably the mineralogy and skeletonized biota.
***_FAC / _FACf	Facies classification in terms of components and features.
(***_DGN)	(A diagnostics file reporting issues found during processing. Comma delimited format suitable for spreadsheet. Not considered further.)
	STAGE 2 OUTPUTS (Merged streams**)**
***_ONE / _ONEf	The EXT, PRS and CLC outputs are telescoped, with precedence to either EXT or PRS (user selectable). This is the primary form of output table that is used in GIS mappings, in concert with CMP, FAC and SRC.
***_ALL / _ALLf	This is simply a sequential re-listing of the EXT, PRS and CLC file contents, under just one column-header record. The EXT, PRS, CLC outputs are concatenated, represented equally. This file is very large, rather unwieldy for GIS packages to handle over large regions.
	STAGE 3 OUTPUTS (Fully integrated**)**

***_WWD / WWDf	This is the ONE file, but with water depths added from a gridded bathymetry such as ETOPO2. Optionally, only (a) the missing or (b) all the water depths, can be placed.
***_FNLf	This is the data for the 'top-20' parameters in its final condition for use in all projects. *This file format is the most highly useable of all.*
***_POS	Statistics on the closeness of fit between input sample water depths and the grid water depths, useful when investigating locational errors.
	STAGE 3 OUTPUTS (Specialized Projects**)**
***_IDC.txt, ***_ISO.txt, ***.WRL, ***_DEC.txt, ***/_RKW, ***/_VRML, ***/LOGS, ***/_TABLS, ***/_CRLYZR, ***/_ADDONS	NOT DESCRIBED FURTHER HERE. Stratigraphic projects involving Corelyser, Rockware, GIS and VRML.
***CEL.kml, ***/_CEL/*.txt, ***/_HTM/*.htm, ***/_BIN.htm, ***_SRC.htm	NOT DESCRIBED FURTHER HERE. Virtual Globe projects of binned surficial sediment data
	STAGE 4 OUTPUTS (Preparations for Gridding**)**
***_CLRf	Compositional Analysis (CoDA) log-ratio transformed values for gravel:sand:mud.

The Parameter Fields

The following lists define the individual output fields, their formats, units, uses, precautions, and known issues.

Source Table Fields

The SRC table is generated for each Data Collection. It is a catalog of the datasets' origins, metadata, keys, extents and entry methods.

PARAMETER	UNITS, MEANING,	COMMENTS
DataSetKey	Unique sequential numeric key to SRC file	For relational linking
SourceCode	Reference name for Source DataSet {Data Collection Code}	Name is from the SRC line in the foundational data (Data Resource Files); the Data Collection Code gives the name of the DRF that contains the data.
DataOwner	Institution or Person who owns the data	Institution name or web uri.
DataPerson	Person facilitating supply of data from Source	Name of the providing person. May include email address. Their institution may differ from the institution with responsibility for the data.
SourceSecur	Level of Confidentiality applied when data was obtained for dbSEABED.	Other measures may also be required to maintain agreements, do not rely completely on this entry.
ReleaseSecur	Level of Confidentiality required on releases of data in its native form.	Other measures may also be required to maintain agreements, do not rely completely on this entry. Other levels may apply to the data once processed in dbSEABED.
AttrEntryMethd	How the seafloor attribute information was entered.	For example: from a data table, from text, or by digitizing a map.
LocationEntryMethd	How the positional information (Lat, Lon) was entered.	For example: from a data table, from text, or by digitizing a map.
DataSourceType	The nature / origin of the document or data	For example: table of analyses, PhD thesis, Published paper, Unpublished report.
LocationRegion	The location and span of the data collecting.	Some indication of the sea/ocean, bay/gulf, etc., country, hemisphere, etc. Free text.

SurveyDate	Date of sampling	May be incomplete or approximate, e.g., “??-08-1995”, or “23-Oct-197?”, or “AustralWinter 2007”
ReportDate	Date on source of dataset (report, digital file, etc)	May be incomplete or approximate, e.g., “??-08-1995”, “23-Oct-197?”, or “AustralWinter 2007”
NavMethod	Free text information on the navigational methods and accuracies.	Navigation technology, model or method; data may include an accuracy estimate.
LocnKey(Start)	First relational Site Key assigned to data in the data set	Alphanumeric, eg: “144567”. Helps to track data items when there is a query.
ObsvnKey(Start)	First relational Observation Key assigned to data (observation, sample, subsample) in the data set	Alphanumeric, eg: “144567”. Helps to track data items when there is a query.
LocnsOutput	Count of the Sites going to output from this dataset	May not correspond to the difference in Keys between this dataset and the next because some keys don’t produce output (fail).
ObsvnOutput	Count of the Observations (observations, samples, subsamples) going to output from this dataset	May not correspond to the difference in Keys between this dataset and the next because some keys don’t produce output (fail QC).
WestBounding	Westernmost coordinate of the data output for this dataset.	Degrees
EastBounding	Easternmost coordinate of the data output for this dataset.	Degrees
NorthBounding	North-most coordinate of the data output for this dataset.	Degrees
SouthBounding	South-most coordinate of the data output for this dataset.	Degrees

SRC example: 24,"Chen++1999_Meiofauna_StraitsMagellan_BeagleChannel {am9_chle}","G. T. Chen","University of Gent","open","open","edittable","edittable","publpaper","Straits of Magellan & Beagle Channel","","1999","", 64226, 88095, 20, 20,-70.94167,-66.89333,-52.995,-55.14

Primary Table Fields

The following arrangement of fields is found in the EXTf, PRSf, CLCf, ONEf, ALLf, WWDf and FNLf files and in some HTML renditions of the same.

PARAMETER	UNITS, MEANING, RANGE	COMMENTS
Latitude	Degrees, WGS 84 Spheroid, 90° to -90° range	WGS 84 Spheroid is within 1m of the more recent International Earth Rotation Service Terrestrial Reference Frame (ITRF) (GDA for Australia)
Longitude	Degrees, WGS 84 Spheroid, -180° to 180° range	
WaterDepth	Metres below Mean Low Sea Level	Not always tidally or (sonar) sound-speed corrected
SampleTop	Metres below seabed surface	
SampleBase	Metres below seabed surface	If Null and Top <> Null then equals Top
SiteName	Survey or laboratory code for site	Not Unique
DataSetKey	Unique sequential alpha-numeric key to SRC file	Use for relational links
SiteKey	Unique sequential alpha - numeric key to SRC file	Use for relational links
SampleKey	Unique sequential alpha-numeric key to SRC file	Use for relational links
Sampler	Type of sampling device (or inspection / probe)	Two parts: (a) simplified code generated in dbSEABED; (b) name as given in the original input data. Note: many cases have no data on sampler type (then marked “UnidDevice”)
DataTypes	For audit only	Indicates type of data contributing to output. Has form: ".....PP--PP---C------CPC-."
Gravel	Gravel grainsize fraction, percent	Wentworth size scale.
Sand	Sand grainsize fraction, percent	Wentworth size scale.

Mud	Mud grainsize fraction, percent	Wentworth size scale.
Clay	Clay grainsize fraction, percent (also included in Mud)	Is output for EXT stream only since can only be determined only by analysis. Uses the science definition of textural clay: <2um size though that standard is hard to apply across all data.
Grainsize	Phi characteristic grainsize	Consensus of mean and median grainsizes
Sorting	Phi grainsize dispersion	Standard deviation sorting only
SeafloorClass	Class (Facies) with the maximum Fuzzy Membership value > 30%	Output for PRS table only
ClassMbrshp	Fuzzy membership (%) of above Class (Facies)	Output for PRS table only
FolkCode	Hydrographic Bottom Type (HBT)e	Refer to Hydrographic Office (1991) for Codes or Coastguard Survey Codes the codings. The EXT output is an echo of naval HBT codes held in the database. The PRS output is an HBT rendition of the textures and grain compositions of all descriptions in the database. The CLC outputs are HBT renditions of the textural (R:G:S:S:C) and weed makeup of the sediments (eg, from numeric grainsize analysis data).
RockMbrshp	Fuzzy membership (%) reflecting percent exposure	Refer to publications.
VegetationMbrshp	Fuzzy membership (%) reflecting percent coverage	Includes seagrasses, kelp and other algae
Carbonate	Percent	Depending on analysis method may or may not include dolomite.
MunsellCode	Standard Alphanumeric coding of color partitioned into Hue, Value and Chroma	Example “5YR 6/4”; refer to Geological Society of America (2009).
OrganicCarbon	Percent	Minimum value from descriptions (PRS tables) is 0.1%
ShearStrength	Log10 of undrained shear strength, in KiloPascals (kPa)	From a variety of instrumentation

Porosity	Percent	Integrates data from porosity, void space, dry and (estimates calculated from) wet bulk densities.
P-WaveVelocity	Metres / second	Usually not corrected for P/T or frequency effects. Most data from atmospheric P/T and high frequencies.
BottomRoughness	In a coding that expresses the height and length of the bottom feature with greatest aspect ratio.	A coded output representing the V:H of the roughness element with greatest aspect ratio, values expressed as (rounded) integer log2 of V and H in cm. For example "4:6" represents 16cm height over length scale of 64cm. Powers <0 are set to zero (i.e., scales <1cm not considered). Horizontal lengths scales are normal to strike and are the length of expression of a feature, rather than wavelength of repetition. The outputs refer only to observed roughness; ephemeral roughness would need to be modeled.
Critical Shear Stress	Log10 of Tau in kPa, being the Shear Stress required to initiate easily observable erosion and transport, whether by traction or suspension.	Taken from a compilation of published relationships ranging from large boulder to muds, through a range of grain shapes (e.g., Shell).
Geologic Age	Millions of years	From the parsing of chronostratigraphic terms, and laboratory various age dating techniques.
PhaseSituation	A string entry that belongs to the Locational data section.	It describes any special circumstance of the sampling or analysis that would cause the result not to refer to the bulk geomaterials.

EXT Example: 60.97361,-2.471845,450,0.00,0.30,"+60-003/369/BC/1",1,10,26,"BoxCorer","TXR",23,74,3,-99,-99.0,-99.0,"",-99,"",-99,-99,-99,"",-99.0,-99.00,-99,-99,"",-99.00,"","-"
PRS Example: 54.04382,-5.715384,37,0.00,0.00,"+54-006/78/GS/1",1,6,17,"Grab:Shipek","SFT",0,77,23,-99,2.5,1.1,"",-99,"mS",-99,-99,-99,"",-99.0,-99.00,-99,-99,"0:0",-99.00,"","surface"
CLC Example: 3.158333,103.91830,27,0.00,0.00,"1958 ; 367(1)",7,84374,157023,"Sampler:Suction","SFT",-99,-99,-99,-99,-0.5,1.1,"",-99,"sG",0,-99,-99,"",-99,-99.00,43,1765,"0:0",-0.02,"","surface"

ONE Example: -51.75450,-57.96750,0,0.00,0.00,"Beach Point",2,6,6,"UnidDevice:| |",".....---------P-------P--.",-99,-99,-99,-99,-99.0,-99.0,"-",-99,"-",100,-99,-99,"-",-99.0,-99.00,-99,-99,"8:12",-99.00,-,"surface"

Audit Formats of EXT, PRS and CLC Tables

The 3 streams have a simple Audit Code which repeats the types of data lines which their data came from. For example with EXT: “TXR” if that was the case for gravel:sand:mud, etc.; “CMP” for carbonate, organic carbon; “GTC” for shear strength, etc. For “PRS” lines: “SFT” for values derived from a “SeaFloorType” word description. An investigator would need to delve back into the raw (foundation) data to ascertain the exact data origins.

Special Audit Formats of ONE, ALL, WWD and FNL Tables

The "ONE" output tables are a telescoping of the EXT, PRS and CLC results (see below) into one table that takes the best quality data of each field, determined field-by-field. The output format is the same as for EXT, PRS and CLC tables except for field 11, which indexes the source of the telescoped data as follows.

Field 11 performs an audit function: where does that data come from ?

In EXT, PRS, CLC files an entry like ". . . . . EEEEEEPPCE-EEE-CC-C." tabulates the origins of the data using 'E' for extracted data, 'P' for parsed data, 'C' for calculated data and '-' for no (null) data. The true data begins at field 12 (gravel) after the position and housekeeping data fields (“.”), and applies to 20 attribute fields.

The Audit Codes do perform very useful functions in the data processing. For instance, uncertainties can be gauged between analysed (E) and descriptive (P) data items. They are also used for tracking real and filled/masked data items in the Compositional Analysis methods.

Note that Field 11 in EXT, PRS and CLC tables performs the same audit function, but at a lower level in the processing of data.

Component/Feature Table Parameters

The CMP table outputs Fuzzy Membership values (%) for each denoted component or feature flagged for inclusion in the project setup file. The list can differ between projects, for instance between Australia (biogenic) and the USA (terrigenous).

PARAMETER	UNITS, MEANING, RANGE	COMMENTS
Latitude	Degrees, WGS 84 Spheroid, 90° to 90° range	WGS 84 Spheroid is within 1m of the more recent International Earth Rotation Service Terrestrial Reference Frame (ITRF) (GDA for Australia)
Longitude	Degrees, WGS 84 Spheroid, -180° to 180° range	
WaterDepth	Metres	Not always tidally corrected

SampleTop	Metres below seabed surface	
SampleBase	Metres below seabed surface	If Null and Top <> Null then equals Top
DataSetKey	Unique sequential numeric key to SRC file	For relational linking
SiteKey	Unique sequential numeric key to SRC file	For relational linking
SampleKey	Unique sequential numeric key to SRC file	For relational linking
RecordType	For audit only	Indicates data status: 1-3 samples, 8 error, 4 time series, etc.
DataTypes	For audit only	Indicates type of data contributing to output. Only indicates the type of data line this information was derived from, e.g., "SFT" for "SeafloorType".

Then, repeated for each component / feature Component Abundance or Feature Intensity:

CmpMbrshp	Fuzzy Membership (as %)	The component or feature must be a major synonym in the dbSEABED dictionary (see list of components below).

CMP Example: -0.44510,9.10960,14.0,0.00,0.00,"(Assigned) 364",4,830,726,"UnidDevice:| |","SFT",-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,33.3,-99,-99,-99,-99,-99,-99,-99,-99,-99,33.3,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,"surface"

The components and features currently output (November 2000 auSEABED settings, designed for work with USGS) are as follows, with abbreviations and details.

In the table header, and also here features are marked with a "_F" postscript on the feature name. Their quantities are intensities, that is, they are not necessarily normalized to 100% and are highly relative and qualitative values.

Abbreviation - Details	**Abbreviation - Details**
fld - feldspar	spng - sponges qtz
- quartz	pectn - pectens sulf -
sulfide	clst - clasts
octcor - octocorals	sil_spic - siliceous spicules
maf - mafics	gstrpd - gastropods
dolmt - dolomite	rlct- - relict materials
metlif - metalliferous	umafic - ultramafics

frm - forams
sftcrl - soft corals
klp - kelp
hvy_min - heavy minerals
lrg_frm - large modern foraminifera (e.g. Marginopora)

wood - wood
crl - coral and octocoral material
crl_dbr - coral debris / material
orgcbn - organic carbon
mica - mica
bfrm - benthic forams
claymin - mineral clay
ool - ooliths and ooids

brach - brachiopods
shl - shells (mollusc & brachiopod)
bryz - bryozoans
burw – burrows
crustac - crustaceans

baslt - basalt
bioturb – all bioturbation
shl_dbr - shell debris / material
ostr – ostracods
gls - glass
brncl – barnacles
vol – volcanics
d alg - algae (hard)

plnk_frm - planktic forams
sol_crl - solitary corals
nan - nannofossils (coccoliths)
silca – hydrous silica
biv - bivalves
ploid – peloids
ptr – pteropods
oyst – oysters
lmp - lumps
gyps - gypsum
cal_nod - calcareous nodules
zeol - zeolites
diat - diatoms
pinna - Pinna (razor clams)
glauc - glauconite
rad - radiolaria
rck_frg - rock fragments
crinod - crinoids
borng - borings
pumc -pumice
trail - trails
trrg - terrigenous
halmda - Halimeda
ophiurd - ophiuroids
fces – faeces
rhodl - rhodoliths
echnd – echinoids
phspht - phosphate
crnalg - coralline algae
weed - 'weed'
fe_nod - ferruginous nodules
srpul - serpulids
seagrs - seagrass
coal - coal

Facies Table Parameters

The FAC table outputs Fuzzy Membership values for each denominated seabed class (facies) for each sample where word-based descriptions are sufficient to support the analysis. The facies are denoted in the setup table "db8_fac.txt" and can be set differently between projects, for instance between Australia (biogenic) and the USA (terrigenous).

PARAMETER	UNITS, MEANING, RANGE	COMMENTS
Latitude	Degrees, WGS 84 Spheroid, 90° to - 90° range	WGS 84 Spheroid is within 1m of the more recent International Earth Rotation Service Terrestrial Reference Frame (ITRF) (GDA for Australia)
Longitude	Degrees, WGS 84 Spheroid, -180° to 180° range	
WaterDepth	Metres	Not always tidally corrected
SampleTop	Metres below seabed surface	
SampleBase	Metres below seabed surface	If Null and Top <> Null then equals Top
DataSetKey	Unique sequential numeric key to SRC file	For relational linking
SiteKey	Unique sequential numeric key to SRC file	For relational linking

SampleKey	Unique sequential numeric key to SRC file	For relational linking
RecordType	For audit only	Indicates data status: 1-3 samples, 8 error, 4 time series, etc.
DataTypes	For audit only	Indicates type of data contributing to output. Only indicates the type of data line this information was derived from, e.g., “SFT” for “SeafloorType”.

Then, repeated for each Facies, the Facies Membership:

Facies	Fuzzy Membership (/1.0) for each type of seabed.	Repeated for each nominated facies as listed in "db8_FAC.txt" setup file (see below).

FAC Example: 53.62900,-4.49500,68,0.00,0.00,"69",189,25228,21311,"UnidDevice:| |","SFT",-99,-99,-99,-99,-99,-99,-99,-99, 31,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,-99,""

Facies Names currently output (November 2000 auSEABED settings, designed for work with USGS) are: CarbMud CarbSand CalcCrust/Nods Oolite Peloid Terrigenous Volcanic 'Coral' 'Shell' CoralReef Sponge Bryozoan CalcPelag SilcPelag Pelagic Phosphate Glauc/Relict Seagrass Plants LargeFrm Rhodolith CorallnAlga Halimeda MnNodule Bioturbated

The defining Components / Features (Senior Synonyms) are: cal_mud mcrt ooz cal_ooz cal_snd skl_dbr limstn calcret carb_nod calcrst ool oolt ooltc ploid trrg hvy_min qtz fld maf vol baslt pumc maf tuff crl octcor srpul crl_dbr crlrf shl mlsc biv gstrpd scph brach crlrf lthmnn spng bryz ooz plnk_frm nan ptr cal_ooz sil_ooz rad silf diat ooz plnk_frm nan ptr rad cal_ooz phspht glauc rlctseagrs plnt weed seawd seagrs klp lrg_frm mrgnpra rhodl algl_nod lthmnn crnalg rhodl halmda mn_nod mn_crst mn_stnd mnoxd mn_mnod burw pit trail fces .

Compositional Analysis Infilling/Masking

Compositional Analysis (CoDA; Aitchison 1986) is required to manipulate and grid gravel:sand:mud ratios data. But in order to do this, the full triplet gravel:sand:mud is required which means that data with only one present (e.g., just mud %) cannot pass to griddings, etc. However, dbSEABED has an inbuilt method of temporarily filling (padding) the gaps with plausible values to compute the log-ratio transforms of the ratios, then removing the pad results before gridding, etc. The transform we use is the CLR method.

The “CLR” files hold the original complete and filled gvl:snd:mud values, the results of the logratio transforms, and the Audit codes that indicate where the padding has taken place and must be undone.

PARAMETER	UNITS, MEANING, RANGE	COMMENTS
Latitude	(As for Primary Files)	
Longitude	“	
WaterDepth	“	
ObsvnTop	“	
ObsvnBot	“	
LocnName	“	

DataSetKey	“	
LocnKey	“	
ObsvnKey	“	
Device	“	
DataTypes	As for primary files, but ‘F’ where a gvl, snd or mud Null is replaced / padded by a Fill value.	The “F” audit codes here indicate where padding has been done and needs to be reversed later
PaddedGravel	Gravel %	But Nulls padded with a value
PaddedSand	Sand %	But Nulls padded with a value
PaddedMud	Mud %	But Nulls padded with a value
PaddedGravelCLR	Log Ratio CLR for gravel fraction	Using the original values where complete, else the padded values
PaddedSandCLR	Log Ratio CLR for gravel fraction	Using the original values where complete, else the padded values
PaddedMudCLR	Log Ratio CLR for gravel fraction	Using the original values where complete, else the padded values
ObsvnDetails	(As for Primary Files)	

Appendix 1. Filetypes by program

FILE TYPE	PRODUCED BY	USED TO GENERATE
***_EXT.txt	dbS_DPR*.bas	***_ONE, ***_ALL, ***_WWDf, & ***_FNLf
***_PRS.txt	"	"
***_CLC.txt	"	"
***_CMP.txt	"	***_CMPf
***_FAC.txt	"	***_FACf
***_SRC.txt	"	*Highly useful as Guide to the Sources of the Data*
(***_DGN.txt)	"	(For run diagnostics only)
***_EXTf.txt; ***_PRSf.txt; ***_CLCf.txt	dbS_INTEGRATE-*.py	(rarely used now)
***_CMPf.txt; ***_FACf.txt		*Most useable for gridding, GIS and RDB projects*
***_ONE.txt	"	(rarely used now; merge to one)
***_ALL.txt	"	(rarely used now; concatenate all)
***_WWDf.txt	"	(rarely used now; is the ONE With Water Depths filled in)
***_FNLf.txt	"	*Most useable for Gridding, GIS and RDB projects*
***_POS.txt	"	(For locational statistics only)
***_CLR.txt	dbS_INTEGRATE-*.py	*CoDA logratios and padding/unpadding data*

Appendix 2. Processing Flow