dbSEABED
Data Mining Output Fields

Contents

Introduction

dbSEABED outputs a set of ASCII tables which can be imported into practically any GIS, Relational Database, Spreadsheet or Maths application.

- The files are comma delimited.
- A header describing the fields (columns) occupies the first line. All Header entries are in quotes ("").
- In the data section all string variables are enclosed in quotes ("") and the data in each column is either all numeric or all string.
- For numerics -99 is the usual 'Default Null' - the marker for 'No Data'; for strings it is "" (no character). An exception is made for Latitudes and Longitudes; the 'Default Null' is -999.
- Files are generated with Intel ('PC') binary arrangements.

The output parameters span a range of themes:

Housekeeping		Data Outputs			
Coordinates	DataOrganization	Textures	Classification	Composition	PhysicalProperties

Top of page

The Output Table Format

Output filenames are named "*_XXX.txt" where * is a prefix set by the operator to denote a study region, for example "syd" for Sydney, "cal" for California. XXX is the type of table. The tables types are as follows.

- **SRC** (*_SRC.txt") describes the dataset source of sites and subsamples.
  Note that records in the EXT, PRS, CLC, ONE and ALL files are relationally keyed to the records of the SRC file using sequential numeric indexes for (i) the dataset, (ii) the site and (iii) the subsample.

- **EXT** ("*_ext.txt") containing data items simply extracted from the data resources through data mining
- **PRS** ("*_prs.txt") containing the results of parsing descriptions in the data resources
- **CLC** ("*_clc.txt") containing results from calculating variables using empirical functions working on the results of extraction or parsing

  EXT, PRS and CLC tables have an identical set of output parameters. This is so they can be combined for GIS operations.
  The 3 sets of data have different reliabilities and clients of dbSEABED may judge for themselves which to use.

  - The EXT data is usually based on instrumental analyses (probe or laboratory) but may apply to just a subsample of the sediment (eg. no large shells).
  - The PRS data is less precise because it comes from word-based descriptions, but will include information on outsized elements, consolidation that are not usually in EXT data.
  - The CLC data is the most derivative and certainly the least accurate; however, many clients appreciate that it extends the coverage of map areas with attributes, especially physical properties attributes.
    Calibrations of the performance of the parsing and calculation processes in dbSEABED are available.

- **ONE** ("*_ONE.txt") files are generated by a program that telescopes the EXT, PS and CLC results in that order of preference.
  This is the primary form of output table that is used in GIS mappings.

- **ALL** (*_ALL.txt") files are generated by a program that concatenates the EXT, PRS and CLC files.
  This file is very large, rather unwieldy for GIS packages to handle over large regions.

- **CMP** ("*_cmp.txt") which lists the Fuzzy Memberships (as percent) of many important features and components of the seabed, for instance of pelagic forams, burrowing, peloids, kelp, etc.

- **FAC** ("*_fac.txt") providing Fuzzy Memberships (as percent) of certain groupings of features and components of the seabed, for instance of manganese nodules, crusts, micronodules and stains in the facies 'MnNodules'.
  The memberships apply to the major synonyms of terms in the data resources. For example 'PlankticForam' is the major synonym of "Globigerina"; and 'Burrow' for "Zoophycos".

- A **2FX** (To Fix, *_2fx.txt") file is also created to report any shortcomings and inconsistencies detected in the data or processing. It is solely for the operators of dbSEABED to use.

Top of page

Source 'SRC' Table Fields

PARAMETER	UNITS, MEANING, RANGE	COMMENTS
DataSetKey	Unique sequential numeric key to SRC file	For relational linking
SourceCode	Reference name for Source DataSet	
DataOwner	Institution or Person who owns the data	
DataPerson	Person facilitating supply of data from Source	May include email address.
ReleaseSecur	Level of Confidentiality required on release of data	
SurveyDate	Date of sampling	May be incomplete, eg. ??-08-1995, or ??-Oct-197?
ReportDate	Date on source of dataset (report, digital file, etc)	
NavAccuracy	Degrees	

SRC example:
1 ,"Reynolds++Tagru-Makassar","NGDC","Purchased via Carla Moore","confidential","17-12-1980","??-??-1981"

Top of page

Extracted 'EXT', Parsed 'PRS' & Calculated 'CLC' Table Fields

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 Nul and Top <> Nul then equals Top
SiteName	Survey or laboratory code for site	Not Unique
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
Sampler	Type of sampling device (or inspection / probe)	In the terms of original input data.
DataTypes	For audit only	Indicates type of data contributing to output.
Gravel	Gravel grainsize fraction, percent	
Sand	Sand grainsize fraction, percent	
Mud	Mud grainsize fraction, percent	
Clay	Clay grainsize fraction, percent (also included in Mud)	Output for EXT only since can only be determined only by analysis
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
HBTCode	Hydrographic Bottom Type	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	
WeedMbrshp	Fuzzy membership (%) reflecting percent coverage	Includes seagrasses, kelp and other algae
Carbonate	Percent	
MunsellCode	Standard Alphanumeric coding of colour partitioned into Hue, Value and Chroma	Example "5YR 6/4"; refer to Geological Society of America (199?).
OrganicCarbon	Percent	Minimum value from descriptions (PRS tables) is 0.1%
ShearStrength	Log_{10} of undrained shear strength, in KiloPascals (kPa)	From a variety of instrumentation
Porosity	Percent	
P-WaveVelocity	Metres / second	Usually not corrected for P/T effects
Bottom Roughness	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 (ie 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 (clast, ripple, mound); ephemeral (eg. ripple) roughness will need to be modelled by clients of dbSEABED.
Critical Shear Stress	Log_{10} of Tau in kPa, being the Shear Stress required to	Taken from a compilation of published relationships ranging from large boulder

	initiate easily observable erosion and transport, whether by traction or suspension.	to muds, through a range of grain shapes (eg. shell).
Backscatter	A non-dimensional index of bottom Acoustic Reverberation Strength used by the Royal Australian Navy (RAN); not to be used for scientific or engineering purposes.	Calculated according to an algorithm designed by DSTO (Australia); values appear only in calculated (CLC) output.

EXT Example:
 -2.48330, 117.60000, 40,-99,-99,"USNS Chauvanet 81-1", 1 , 2 , 2 ,1,"T ", 2, 91, 5, 2, +1.8, +2.0,"-",-99,"", 0,-99,-99,"",-99,-99,-99, -99, "2:3",-99,-99
PRS Example:
-16.73700, 115.53500, 2189, 0.00, 62.70,"ODP 761", 3 , 11 , 280 ,1,"L ", 0, 23, 76,-99, +6.4,-99.0,"Pelagic", 77,"",-99,-99,100,"",-99,-99,-99, -99, "0:0",-99,-99
CLC Example:
-35.08200, 109.09500, 5435,-99,-99,"Robert Conrad RC08 055", 6 , 836 , 1135 ,1,"PR:L",-99,-99,-99,-99, +4.8, +2.1,"",-99,"M.S",-99,-99,-99,"",-99,-99, 51,1693, "2:4",-1.5,1..5

Top of page

Telescoped 'ONE' Table Fields

This is the primary form of table used in GIS mappings.

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. An entry like "EEEEEEPPCExEEExCCxCC" tabulates the origins of the data using 'E' for extracted data, 'P' for parsed data, 'C' for calculated data and 'x' for no (null) data. It begins this at field 12 (gravel%) and applies to 20 attribute fields.

Example: -33.17000,+151.60001,20,-99,-99,"1537",1 ,8 ,8 ,1,EEEEEEPPCExEEExCCxCC,0,100,0,0,+3.0,+1.0,"Terrigenous",82,"S",0,-99,7,"5YR6/3",0,-99,42,1868,"2:3",-3.28 ,2

(Note: Field 11 in 'EXT', 'PRS' and 'CLC' tables performs the same audit function, but at a lower level in the processing of data, at the level of extracting data from the core dbSEABED data resource files.)

Concatenated 'ALL' Table Fields

"ALL" output tables are simply a sequential re-listing of the EXT, PRS and CLC file contents, under just one column-header record. The field formats are the same as for those files.

Component Table Parameters

The CMP table outputs Fuzzy Membership values (as %) for each denoted component flagged for inclusion in the dbSEABED dictionary. 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 Nul and Top <> Nul 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.

Component	Fuzzy Membership (as %) for component which must be a major synonym of components or features.	Repeated for each component entry nominated in the dbSEABED dictionary (see below).

CMP Example:
-16.14500,110.29900, 5702, 46.35, 47.55, 2 , 5 , 98 ,3,"P ",,, 1,,,,,,,,,, 0,,,,,,,,,,,, 26,,,, 72,,,

Components and features currently output (November 2000 auSEABED settings):

fld - feldspar
qtz - quartz
sulf - sulfide
maf - mafics
metlif - metalliferous
hvy_min- heavy minerals
mica - mica
claymin - mineral clay
silca - hydrous silica
gyps - gypsum
zeol - zeolites
glauc - glauconite
rck_frg - rock fragments
umafic - ultramafics
baslt - basalt
gls - glass
vol - volcanicsd
pumc -pumice
trrg - terrigenous
phspht - phosphate
fe_nod - ferrugineous nodules
coal - coal

dolmt - dolomite
frm - forams
lrg_frm - 'large foraminifera' (e.g. Marginopora)
bfrm - benthic forams
plnk_frm - planktic forams
nan - nannofossils (coccoliths)
ptr - pteropods
diat - diatoms
rad - radiolaria
spng - sponges
sil_spic - siliceous spicules
shl - shells (mollusc & brachiopod)
shl_dbr - shell debris / material
alg - algae (hard)
halmda - Halimeda
rhodl - rhodoliths
crnalg - coralline algae
srpul - serpulids

octcor - octocorals and octocoral material
sftcrl - soft corals (living)
crl - coral
crl_dbr - coral debris / material
sol_crl - solitary corals
biv - bivalves
oyst - oysters
pinna - Pinna (razor clams)
pectn - pectens
gstrpd - gastropods
brach - brachiopods
bryz - bryozoans
crustac - crustaceans
ostr - ostracods
brncl - barnacles
crinod - crinoids
ophiurd - ophiuroids
echnd - echinoids
weed - 'weed'
seagrs - seagrass

klp - kelp
wood - wood
orgcbn - organic carbon
ool - ooliths and ooids
ploid - peloids
lmp - lumps
cal_nod - calcareous nodules
clst - clasts
rlct- - relict materials
burw - burrows
bioturb - all bioturbation
borng - borings
trail - trails
fces - faeces

[To Dictionary](#)

[Top of page](#)

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 Nul and Top $<>$ Nul 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.
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:
-12.947 , 117.893 , 5667 , 97.67, 98.00, 2 , 6 , 169 ,,,,,,,0.51,,,,,,,0.04,0.08,0.09,,,,,,,,,,

Facies currently output (November 2000 auSEABED settings):

Facies Name	Defining Components / Features (Senior Synonyms)
CarbMud	cal_mud mcrt ooz cal_ooz
CarbSand	cal_snd skl_dbr
CalcCrust/Nods	limstn calcret carb_nod calcrst
Oolite	ool oolt ooltc
Peloid	ploid
Terrigenous	trrg hvy_min qtz fld maf
Volcanic	vol baslt pumc maf tuff
'Coral'	crl octcor srpul crl_dbr crlrf
'Shell'	shl mlsc biv gstrpd scph brach
CoralReef	crlrf lthmnn
Sponge	spng
Bryozoan	bryz
CalcPelag	ooz plnk_frm nan ptr cal_ooz
SilcPelag	sil_ooz rad silf diat
Pelagic	ooz plnk_frm nan ptr rad cal_ooz
Phosphate	phspht
Glauc/Relict	glauc rlct-
Seagrass	seagrs
Plants	plnt weed seawd seagrs klp
LargeFrm	lrg_frm mrgnpra
Rhodolith	rhodl algl_nod
CorallnAlga	lthmnn crnalg rhodl
Halimeda	halmda
MnNodule	mn_nod mn_crst mn_stnd mnoxd mn_mnod
Bioturbated	burw pit trail fces

[Top of page](#)

Chris Jenkins ([Email](#))
INSTAAR, University of Colorado
23 Sep 2003