

Regulation Documentation Update September 23, 2014

General Description:

This database contains federal commercial groundfish regulations from 1983-2013. Additional years will be added as they become available. This database contains three types of data: regulations governing trip limits; Rockfish Conservation Area (RCA) data; and OY, ABCs for groundfish species and species groups managed by the Pacific Fisheries Management Council. This database is NOT to be used for law enforcement purposes - its sole purpose is for support of research and stock assessments.

The data are stored in a series of SQL Server tables in the CALCOM database. These data come from Federal Registers and from Documents provided by John Devore (PSMFC, Portland, OR). The tables use a technology called FullText Indexing to search large text fields.

Obtaining data from the tables can be done in a variety of ways. The simplest way for most users will be via the CALCOM website. The majority of data retrievals can be made this way and the data can be downloaded to text files by the user. For those with good SQL skills, they can connect directly to the database using a variety of software packages including: R, Excel, and other ODBC compliant software packages. Finally, for special requests, the user can contact Don Pearson and request a data retrieval.

Structure of the data and general concepts:

Tables:

REG_COMMERCIAL_REGULATIONS

This table contains the actual regulations. There is a fulltext index on the regulations column and the global thesaurus can be used on that column. All species are listed in the regulation column by common name or management group (for example: bocaccio or sebastes complex). Due to software issues, the "other fish" and "other flatfish" groups have been renamed to "oth fish" and "oth flatfish". This table can be joined to the REG_ALL_LOCATIONS table via the location_id column.

Column Name	Data Type	Nulls	Comments
UID_Pk	Integer	N	auto generated
Regulation_Date	Date	N	
Location_Id	NVarchar(200)	N	
Regulation	NVarchar(2000)	N	

REG_ALL_LOCATIONS

This table contains points on the coast used for regulations. For example, some regulations refer to Latitude 40 degrees 10 minutes and northward. Each location_id has an associated state, INPFC area, minimum and maximum latitude. As a result, this table can be used to extract regulations pertaining to any part of the coast desired.

Column Name	Data Type	Nulls	Comments
Location_Pk	Integer	N	auto generated

Location_Id	Nvarchar(200)	N	
North_Lat	Real	N	ddmm.mm
South_lat	Real	N	ddmm.mm
States	nvarchar(200)	N	complete name
INPFCs	nvarchar(100)	N	complete name

REG_OY_ABC_OFL

This table contains the various OYs, ABCs, etc for all species and management groups from 1983 through 2013. It is important to note that the definition of ABC changed in 2011 and therefore ABCs are presented in 2 columns, ABC_pre2011 and ABC_post2010. The species or management groups are the same as used in the REG_COMMERCIAL_REGULATIONS table. The location_id column can be matched to the location_id column in the REG_ALL_LOCATIONS table. It is assumed that users of this table are familiar with all concepts relating to the values presented.

Column Name	Data Type	Nulls	Comments
Policy_id	Integer	N	auto generated
Year	Integer	N	
Location_id	nvarchar(50)	N	
OY	Real	Y	
ABC_Pre2011	Real	Y	
ABC_Post2010	Real	Y	
OFL	Real	Y	
ACL	Real	Y	
HG	Real	Y	
Comments	nvarchar(500)	Y	

REG_RCAS

This table contains the locations and timing of the Rockfish Conservation Areas (RCAs) for California, Oregon, and Washington. The area column refers to latitude range. The depths are in the form of nnn - nnn where nnn is a depth in fathoms for the restriction.

Column Name	Data Type	Nulls	Comments
Year	Integer	N	
Month	Char(3)	N	
Area	nvarchar(50)	N	
Location_id	nvarchar(50)	N	for linking to REG_ALL_LOCATIONS
Depths	nvarchar(10)	N	
Gear_type	nvarchar(15)	N	Trawl or nontrawl

Comments nvarchar(200) Y

REG_SPECIES_GROUPS

This table is used by the website for generating queries. It has no enduser value.

Stored Procedures:

REG_LOCATIONS_SP: This stored procedure returns all regulations for all species between a specified northern latitude and southern latitude.

Input Parameters: north latitude, south latitude as ddm.mm

Values returned: regulation_date, text of the regulation, north and south latitude for the specific regulation.

REG_SPECIES_OY_SP: This stored procedure selects all columns from the REG_OY_ABC_OFL table for a specified species.

Input Parameters: species as common name, species_group

Values Returned: species_or_group, Year, OY, ABC_pre2011, ABC_post2010, OFL, ACL, HG, Comments

Note: a species group is always required. A list of the species and groups is in Appendix A of this document.

REG_GRP_WEB: This stored procedure is used by the website to help query species regulations requiring access to the thesaurus.

Web Pages:

Currently there are 7 things that can be done from the CALCOM website:

1. Download documentation
2. Retrieve species specific regulations
3. Retrieve species and management group regulations
4. Retrieve all regulations in effect within a specified range of latitudes
5. Retrieve all regulations pertaining to either Open Access or Limited Entry
6. Retrieve all RCAs
7. Retrieve OYs, ABCs, OFLs, and HGs for a species

In all cases the data can be downloaded to a text file with the “|” used as a column delimiter.

REG_SPECIES_REGS_SP: This stored procedure returns all regulations for a specified species or group.

Input Parameters: species as common name, species_group

Values Returned: regulation_date, text of regulation, northern and southern latitude for the regulation.

Note: a species group is always required. A list of the species and groups is in Appendix A of this document.

REG_TIME_SP: This procedure returns all regulations for all species for a specified time interval.

Input Parameters: start year as yyyy, end year as yyyy

Values returned: Regulation_date, Location_id, text of the regulation

Examining Your Results

The nature of the regulations is somewhat complicated. Often a regulation will pertain to more than one species. For example, if you want data for sablefish, you retrieval will return data for deepwater complex. Occasionally, a species may be listed as part of a group but still have separate regulations and so the group and the species regulations will both be returned. The user should examine the results, sorting by date and location id and bearing in mind that there are different management groupings like open access, midwater trawl etc... It is up to the user to make sure that their retrievals take these issues into consideration.

Accessing the Data

The regulations tables reside in the CALCOM database which exists on the PORICHTHYS server at the Santa Cruz Laboratory. The database is readily accessible to anyone through the use of ODBC connections. In addition, simple queries for the regulations can be accessed through the CALCOM website. This documentation package includes examples of how to write SQL queries to access the data using the fulltext indexes and the thesaurus. In addition, I have created several stored procedures which can be used from within various software packages to access the data.

QUERY EXAMPLES

Simple query for all sablefish regulations

```
SELECT *
FROM reg_commercial_regulations
WHERE CONTAINS(regulation,'sablefish')
```

This query will get all regulations for DTS or DWC including individual regulations for sablefish, thornyheads, and Dover sole. Note in the example DTS is used; however, DWC could be used and the same results would be obtained since they are synonyms.

```
SELECT *
FROM reg_commercial_regulations
WHERE CONTAINS(regulation,'FORMSOF(THESAURUS, sablefish)')
```

This query will return all regulations that pertain to rosy rockfish. Since there are no species specific regulations for rosy rockfish, it returns Sebastes Complex regulations from 1983-1999 and the shelf rockfish regulations from 2000 on.

```
SELECT *
FROM reg_commercial_regulations
WHERE CONTAINS(regulation,'FORMSOF(THESAURUS, rosy)')
```

This query will return all regulations that pertain to chilipepper rockfish. Since there are no species specific regulations for chilipepper rockfish prior to 1999, it returns Sebastes Complex regulations from 1983-1998, then it returns both chilipepper and sebastes complex in 1999, and then from 2000 onwards it returns chilipepper regulations for the south and shelf rockfish regulations for the north.

```
SELECT *
FROM reg_commercial_regulations
WHERE CONTAINS(regulation,'FORMSOF(THESAURUS, chilipepper)') or
CONTAINS(regulation,'chilipepper')
```

This query is used to get the oth flatfish management group (termed Other flatfish in the regulations). Note that it will not get individually named species unless they are explicitly included as part of the regulation which contains the search term 'oth flatfish'.

```
SELECT *
FROM reg_commercial_regulations
WHERE CONTAINS(regulation,'oth AND flatfish')
```

This query is used to get the oth flatfish management group (termed Other flatfish in the regulations) as well as all other flatfish since the thesaurus has synonyms of sole, flounder, and turbot for the search term of flatfish. Note that it will not get individually named species unless they are explicitly included as part of the regulation which contains the search term 'oth flatfish'.

```
SELECT *
FROM reg_commercial_regulations
WHERE CONTAINS(regulation,'oth AND flatfish') and CONTAINS(regulation,
'FORMSOF(THESAURUS,flatfish)')
```

This query is used to get the regulations for sablefish in California. Since some regulations cross the border into Oregon and even Washington, those regulations are also returned.

```
SELECT a.regulation_date, a.regulation, b.location_id, b.states
FROM reg_commercial_regulations a, reg_all_locations b
WHERE CONTAINS(a.regulation,'sablefish') and a.location_id=b.location_id and
CONTAINS(b.states, 'California')
```

This query looks for Open Access regulations.

```
SELECT *
FROM reg_commercial_regulations
WHERE CONTAINS(regulation, 'open AND access')
```

NOTES ABOUT QUERIES:

1. The words "AND" and "OR" are logical operators and therefore are reserved. A query that tried to search for all occurrences of "AND" in the regulations would not return any values.
2. There are a large number of "STOP" words that will not return any values. Words like "other", "some", and others are explicitly ignored by the FULLTEXT INDEX.
3. INPFCs and States are spelled out, not abbreviated eg... California or Conception not CA or CP.
4. Queries are never case sensitive.
5. Location_ids are informative, here are some examples:
 - 4010 North means Latitude 40 degrees 10 minutes north to Canada
 - 3427 3600 – meansBetween latitude 34 degrees 27 minutes (Point Conception) and latitude 36 degrees 0 minutes
 - CONCEPTION – means the Conception INPFC area

Using Stored Procedures

Stored procedures are easy to use. From any place that accepts a query run the following:

```
EXECUTE <stored procedure name> <parameter>, <parameter - if required>
```

Example: EXECUTE reg_time_sp 1998, 2003

Appendix A - List of all species and associated groups in the system

Notes:

1. For some queries and stored procedures, you must include both a species and a group for the query to retrieve all of the data. The retrievals are not case sensitive and plural forms are acceptable eg... sanddab and sanddabs are treated the same.
2. Rockfish are a special case since they can be managed as a distinct species or part of a group. In addition, the groups have changed over time. The fulltext index is “smart” enough to deal with this as long as the thesaurus is used in the retrieval. Any unlisted rockfish can be used in query as long as the thesaurus is used in the query – be sure to use the common name.

SPECIES	GROUP
Rex sole	Oth Flatfish
sanddab	Oth Flatfish
dover sole	DWC or DTS
english sole	Oth Flatfish
petrale sole	Oth Flatfish
arrowtooth flounder	Oth Flatfish
halibut	Halibut
starry flounder	Oth Flatfish
cabezon	Oth fish
scorpionfish	Oth fish
whiting	whiting
lingcod	lingcod
sablefish	DWC or DTS
longspine thornyhead	DWC or DTS
shortspine thornyhead	DWC or DTS
thornyhead	DWC or DTS
Rougheye	Sebastes Complex
Pacific Ocean Perch	Sebastes Complex
Black	Sebastes Complex
Blue	Sebastes Complex
chilipepper	Sebastes Complex
bocaccio	Sebastes Complex
cowcod	Sebastes Complex
yellowtail	Sebastes Complex
yelloweye	Sebastes Complex
yellowmouth	Sebastes Complex
splitnose	Sebastes Complex
bronzespotted	Sebastes Complex
widow	Sebastes Complex
darkblotched	Sebastes Complex
bank	Sebastes Complex
canary	Sebastes Complex

vermilion	Sebastes Complex
blackgill	Sebastes Complex
shortbelly	Sebastes Complex
sebastes complex	Sebastes Complex
nearshore rockfish	Sebastes Complex
shelf rockfish	Sebastes Complex
slope rockfish	Sebastes Complex
oth rockfish	Sebastes Complex
remaining rockfish	Sebastes Complex
rockfish	Sebastes Complex
dogfish	Oth fish
longnose skate	Oth fish
DTS	DWC or DTS
DWC	DWC or DTS
oth flatfish	Oth Flatfish
oth fish	Oth fish

Appendix B

Retrieving Data From Fulltext Indexed Tables

This section briefly describes how to use fulltext indexed tables in SQL Server. It is not a full description, it is simply a guide to the basic concepts. This section does not describe how to create a fulltext indexed table.

There are two basic ways of using fulltext indexing: FREETEXT and CONTAINS. Both methods are used in the WHERE clause of the query. These approaches can use a user defined Thesaurus which must be explicitly called by the query. If the thesaurus is not called, the queries will not use substitution of words although they will use forms of the word. An important concept to understand is that some words are reserved (referred to as "stop" or "noise" words) and these should not be used as part of the search.

The following descriptions are very simple versions. These are very powerful and have many variations in their usage

METHOD 1: FREETEXT

I have not used this method and do not see any particular advantage at this time for my purposes.

This method searches for the *meaning* of the search term, not just a match for the term.

Syntax:

```
SELECT *
FROM <tablename>
WHERE FREETEXT(<column name>, '<search term>')
```

Example:

```
SELECT *
FROM reg_commercial_regulations
WHERE FREETEXT(regulation, 'he appreciated')
```

In the above example, records would be chosen that matched the meaning of 'he appreciated' in the regulation column. For example, records with 'he liked', or 'he respected' might be returned. Note that these similar terms do not need to be included in a thesaurus.

METHOD 2: CONTAINS

This is the approach that I have focused on for my current work.

This method searches for matches for the search term or variants but does not attempt to search for the meaning of the search term.

Syntax:

```
SELECT *
FROM <tablename>
WHERE CONTAINS(<column name>, '<search term>')
```

Examples:

```
SELECT *
```

```
FROM reg_commercial_regulations
WHERE CONTAINS(regulation, 'rockfish')
```

In the above example, the column regulation would be searched for the word rockfish and variants (for example: rockfishes) would be returned.

```
SELECT *
FROM reg_commercial_regulations
WHERE CONTAINS(regulation, 'FORMSOF(THESAURUS, dts)')
```

In the above example, the thesaurus is called and the results returned would be any defined synonyms of the term 'dts'. The thesaurus is described below.

In both methods, the queries can be complex. Multiple tables can be included, even from different databases. Additional filters can be included in the where clause using AND, OR, NOT, etc... . In addition, multiple columns can be included in the fulltext search.

This methodology is powerful enough, that aspects like degree of similarity and proximity of terms can be included to help with the searches. I have not included these features in this note for brevity but it is important to realize that this topic is large and the user may wish to explore other facets of the retrievals.

Appendix C About the Thesaurus

Thesaurus Entry Types:

```
<replacement>  
<pat> bocaccio </pat>  
<sub> shelf rockfish </sub>  
<sub> paucispinus </sub>  
</replacement>
```

In the above example, bocaccio would not be returned, instead, whenever, bocaccio was used in the search term, shelf rockfish and paucispinus would be returned because the <pat></pat> is specifying a replacement in the search term.

```
<replacement>  
<sub> bocaccio </sub>  
<sub> shelf rockfish </sub>  
<sub> paucispinus</sub>  
</replacement>
```