Lummi Intertidal Baseline Inventory

Appendix J:
Relational Database Applications

Prepared by:

Lummi Natural Resources Department (LNR)
2616 Kwina Rd.
Bellingham, WA 98226

Contributors:

Craig Dolphin	LNR Fisheries Shellfish Biologist
Michael LeMoine	LNR Fisheries Habitat Biologist
Jeremy Freimund	LNR Water Resources Manager

March 2010

This page intentionally left blank.

Executive Summary

Data obtained during the LIBI study were stored and analyzed using five custom-developed Microsoft Access 2000 databases that are included on the LIBI DVD. A brief overview of the table structure and user interface of these database applications is presented in this appendix.

Using relational databases to store and analyze the data allowed for rapid data analysis and reporting of the results, as well as enabling interoperability with analytical software such as Geographic Information System mapping software and statistical analysis software.

Table of Contents

1.0 Introduction

One of the goals of the Lummi Intertidal Baseline Inventory is to provide survey data results in a format that is easily accessible for all parties who may have a need to access the data directly. This will be particularly useful should questions arise about taxa that were not considered separately in this study. It will also enable the data to be analyzed based on geographical sub-areas that may be defined differently in the future. For example, in the event of an oil spill, the oil-affected area could be mapped and the data from this study analyzed separately for that specific area.

Although spreadsheets could have been used to achieve this goal, spreadsheets are not the optimum solution for storing and querying large data sets because they lack input validation constraints. As a result, the content of spreadsheets is inherently less predictable than other data storage solutions such as relational databases that enforce validation rules. The choice of data storage solutions has important ramifications when interoperability with other software is important. In this case, interoperability was necessary to use the collected data with mapping software and with statistical analysis software programs. Relational databases also provide more efficient ways to query and reorganize data that can then be exported for analysis in a spreadsheet software program.

Accordingly, a relational database application was developed for each of the four major survey methods used in the LIBI and are included on the attached LIBI DVD. The purpose of these databases is to provide the raw data used to generate the final report of the Lummi Intertidal Baseline Inventory at the same level of detail that was collected in the field. In addition, the four primary databases export data to a fifth access database that is used to provide summary results that are used in related ArcGIS 9.3 (ESRI) mapping projects as described in Appendix K.

The purpose of this appendix is to provide a brief overview of these databases to assist interested parties in exploring the information that is stored in them. It is assumed that the reader has a working knowledge of the use of relational database systems, in particular Microsoft Access 2000.

2.0 Methods

All database applications were developed using Microsoft Access 2000 database software. Form controls and parameters that were used to add, delete, or edit data values have been disabled or hidden in the released version of the databases to prevent accidental modification of the data. However, all code and controls are still present and can be re-enabled from the form's design view if required. Neither the data nor the databases have been protected with passwords or user-level security to ensure that the data can be easily accessed in the future.

Wherever possible, standard data normalization rules were followed to ensure storage efficiency and to facilitate easy query building.

3.0 Results

3.1 Benthic Survey Database

3.1.1 Tables Summary

Data from the dig survey and the epibenthic rock survey of benthic invertebrates and plants are located in the DigSurveyDB.mdb database application. Figure J.1 shows the tables present in the DigSurveyDB.mdb file, and defines the relationships between the tables.

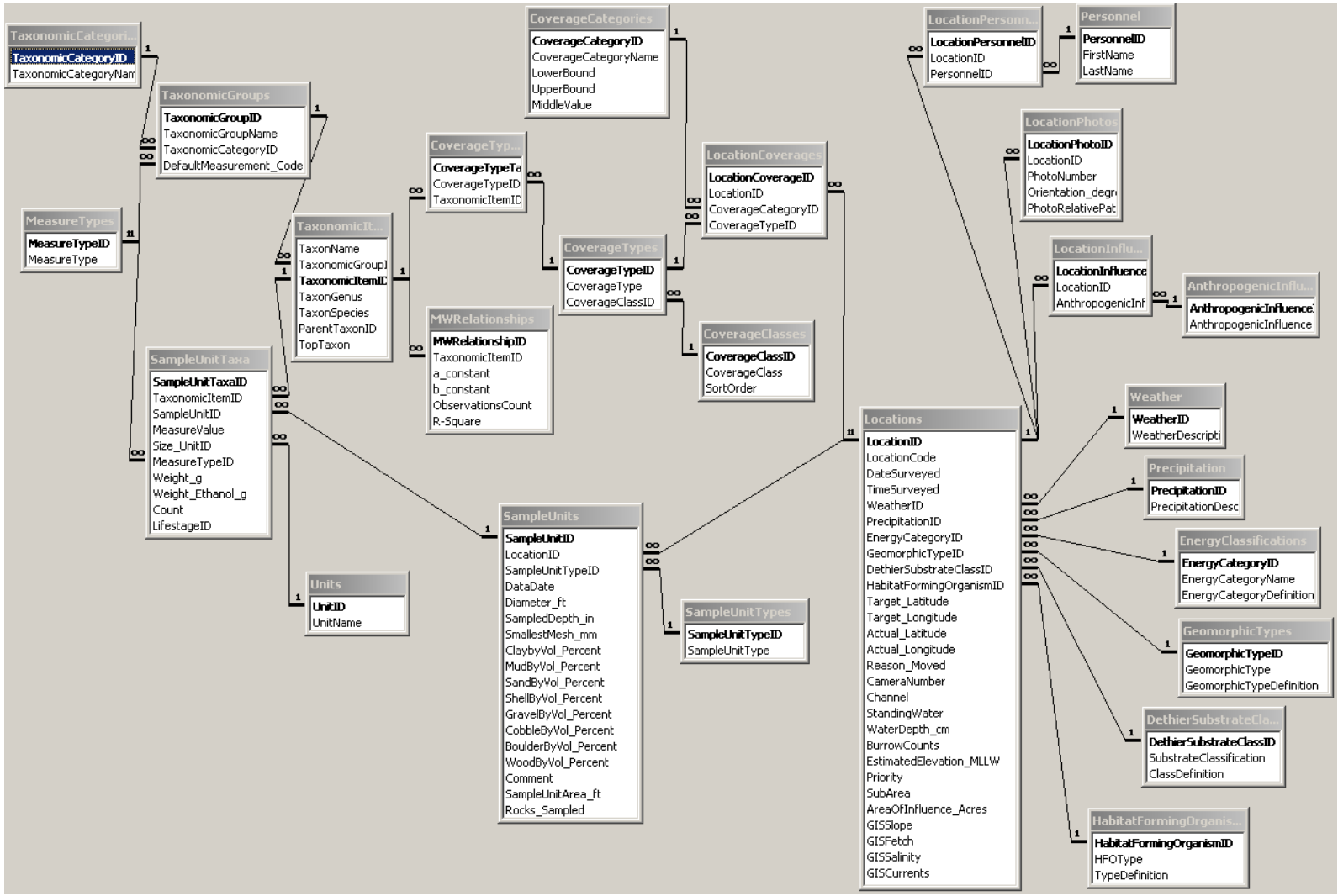

Figure J.1. Entity Relationship Diagram for DigSurveyDB.mdb

While it is not the intent of this appendix to document each table, it should be noted that the three most important tables in this database are the *Locations* table, the *SampleUnits* table, and the *SampleUnitTaxa* table.

The *Locations* table contains information specific to each location that was sampled, including the spatial location (Actual_Latitude and Actual_Longitude) as well as information about various environmental parameters that were measured in the field or that were obtained using various GIS analyses, as described in Appendix H.

There are also several associated tables (e.g., *EnergyClassifications*, *GeomorphicTypes*, *DiethierSubstrateClassifications*) that provide the names and, where appropriate, the definitions associated with the stored ID values for several of these fields.

The *SampleUnits* table stores information that depends on both the location and the type of survey methodology used. For example, the number of rocks sampled is only recorded for the epibenthic rock survey method; and the size of the sampled area is only applicable to the dig survey methodology. This distinction is necessary to ensure that counts of organisms attached to rocks can be excluded from counts of organisms present in the dig survey that might be double-counted if the field crews washed off some of the organisms into the samples after counting them on the rocks.

The *SampleUnitTaxa* table stores information about the organisms found using a particular survey method at a particular survey location. The names of the various taxa that were encountered are stored in the related *TaxonomicItems* table, and the higher-level taxonomic groups and categories are stored in the related tables *TaxonomicGroups* and *TaxonomicCategories* respectively.

The values in the *MWRelationships* table were calculated by the database using the raw survey data as inputs. The table stores species-specific constants (a, b) that are used to calculate weight from size measurements using a power formula where Weight(g) = $a*(measure^{b})$. The code used to calculate these constants can be found in the vba code module *modMiscellaneous* in the public Function *UpdateSizeWeightData()*.

The table *LocationCoverages* stores information about different surface coverages (*CoverageTypes*) that can be found at each location. Where applicable, the table *CoverageTypeTaxa* stores information about any organisms that are typically associated with those coverage types. For example, the Pacific Blue Mussel (*Mytilus trossulus*) is automatically associated with the surface coverage type ‘Mussel Mat’.

3.1.2 User Interface

By default, the database window will be hidden from view when the database opens. This is by design to prevent accidental erasure of data or database objects by inexperienced Access users. Experienced access users can view the tables and other database objects by un-hiding the database window from the Window menu.

Figure J.2 shows the main menu form that appears when the Dig Survey database is opened.

Figure J.2. Dig Survey Database Main Menu

There are seven buttons on the main menu. The two buttons at the bottom of the menu will either close the database altogether (*Stop*), or simply close the menu (*Close*) while leaving the database file open.

3.1.2.1 Viewing taxonomic presence/absence results

The buttons entitled *Location Diversity List* and *Location Diversity Table* both open pre-constructed queries that summarize which taxa were present at each site, along with the site name and spatial location. The first query lists each species on a separate line for each site. The second query creates a summary table that allows for easy comparison of the results. These queries combine the results from the epibenthic rock survey method, the dig survey method, and any known associations with biotic surface coverage types (such as mussel mats or barnacles).

3.1.2.2 Viewing Results for each Site

The topmost button (*Dig Survey Data*) opens up a different form that allows the user to explore the specific results from each location that was surveyed (Figure J.3). The form itself has four tabs that are designed to have strong resemblance to the paper field form design (Appendix A).

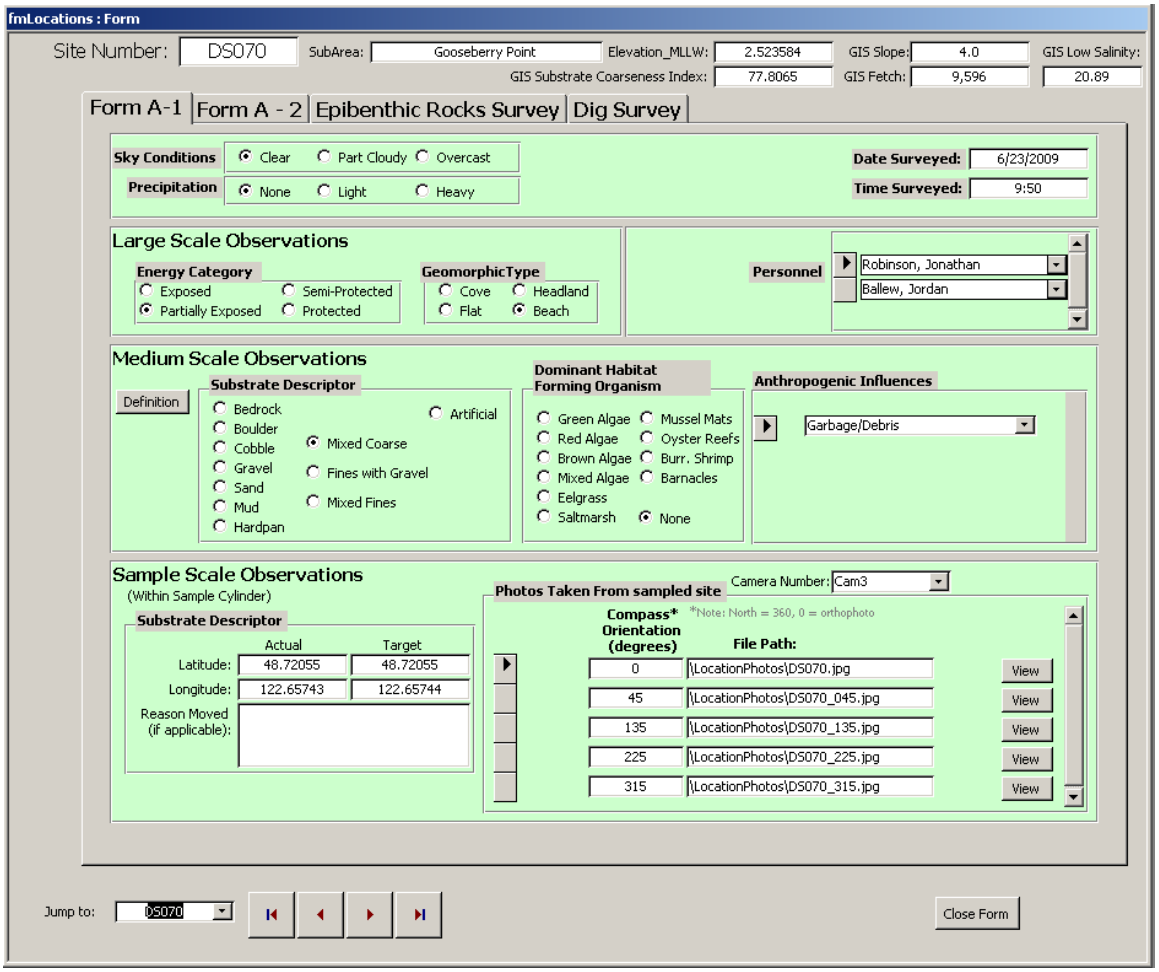

Figure J.3. Locations Form Showing the First Results Tab for Location DS070

The Locations Form is headed with information specific to the site including: the site number, a geographic sub-area descriptor, and statistics derived from various GIS analyses as described in Appendix H.

The user may navigate to results for different sites by using the navigation buttons at the bottom of the form (Figure J.4). The drop-down *jump to* navigation list allows the user to select a specific site, or the user can click on the arrow buttons to view the first site, the previous site, the next site, or the last site in the database.

Figure J.4. Navigation Tools to Change the Site Being Viewed

The first tab (labeled *Form A-1*) shows the information recorded by the field crews relating to habitat descriptors, conditions, and spatial location at the site. Large-scale observations apply to spatial scales of 100 – 1,000 feet. Medium scale observations apply to spatial scales of 10 – 100 feet. Sample scale observations are for the area within the sampling cylinder only.

Photographs taken at each site can be viewed by clicking the *View* button next to the photograph of interest. The compass bearing along which the photograph was taken is indicated by the value in the Compass Orientation field. A value of 0 in this field indicates that the photograph was orthographic (i.e., directed downwards onto the substrate from above as shown in Figure J.5). Values from 1 – 360 indicate that the photograph is taken from the sample location with a bearing equivalent to the degrees of a compass (e.g., Figure J.6).

Figure J.5. Orthographic Photograph of Location DS070 Showing a 6-Inch Reference Item

Figure J.6. A Photograph Taken from Location DS070 with a Bearing of 135 Degrees

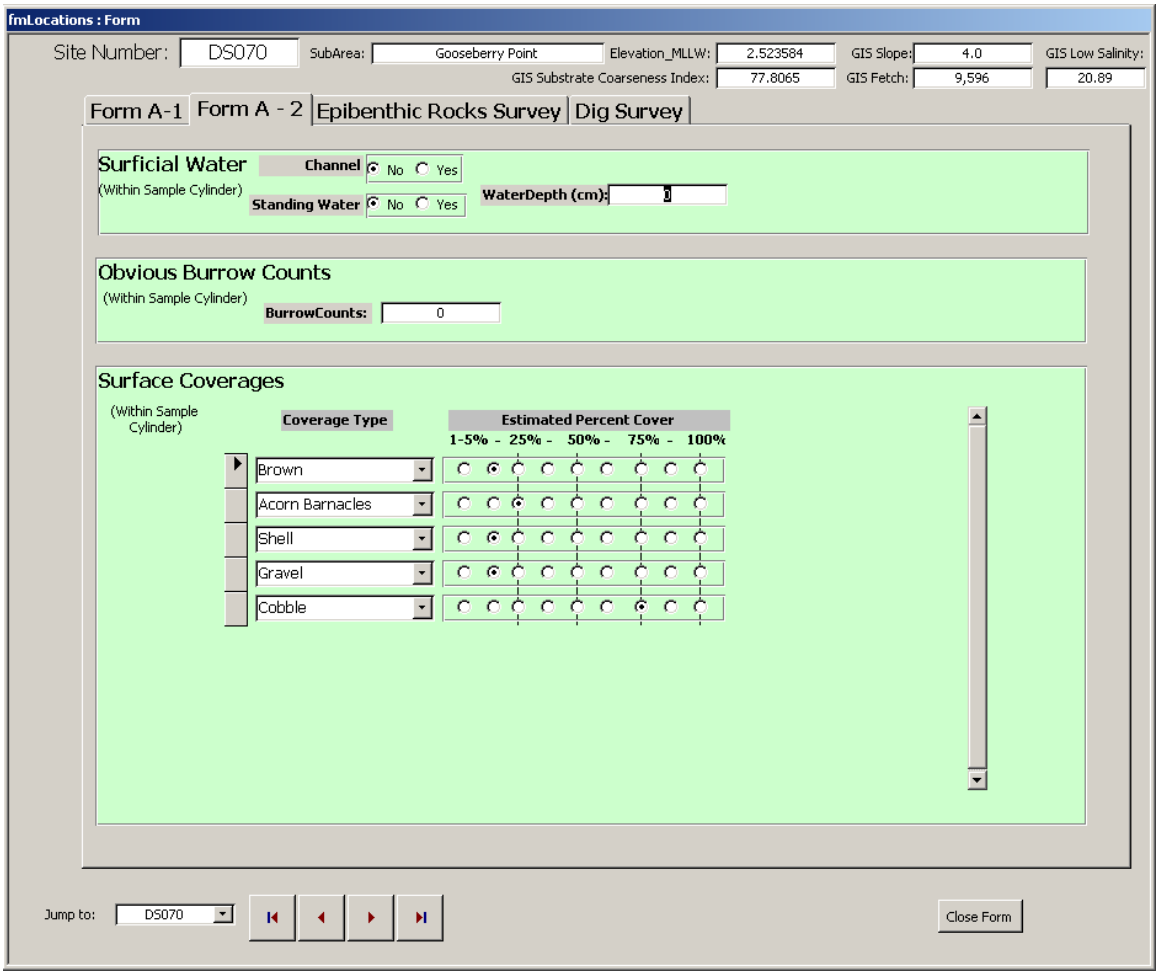

Figure J.7. Locations Form Showing the Second Results Tab for Location DS070

The second tab on the Locations Form shown in Figure J.3 (labeled *Form A-2*) shows the information recorded by the field crews relating to surficial characteristics such as the presence and depth of standing or flowing water, obvious burrows, and surface coverages within the sampling cylinder (Figure J.7).

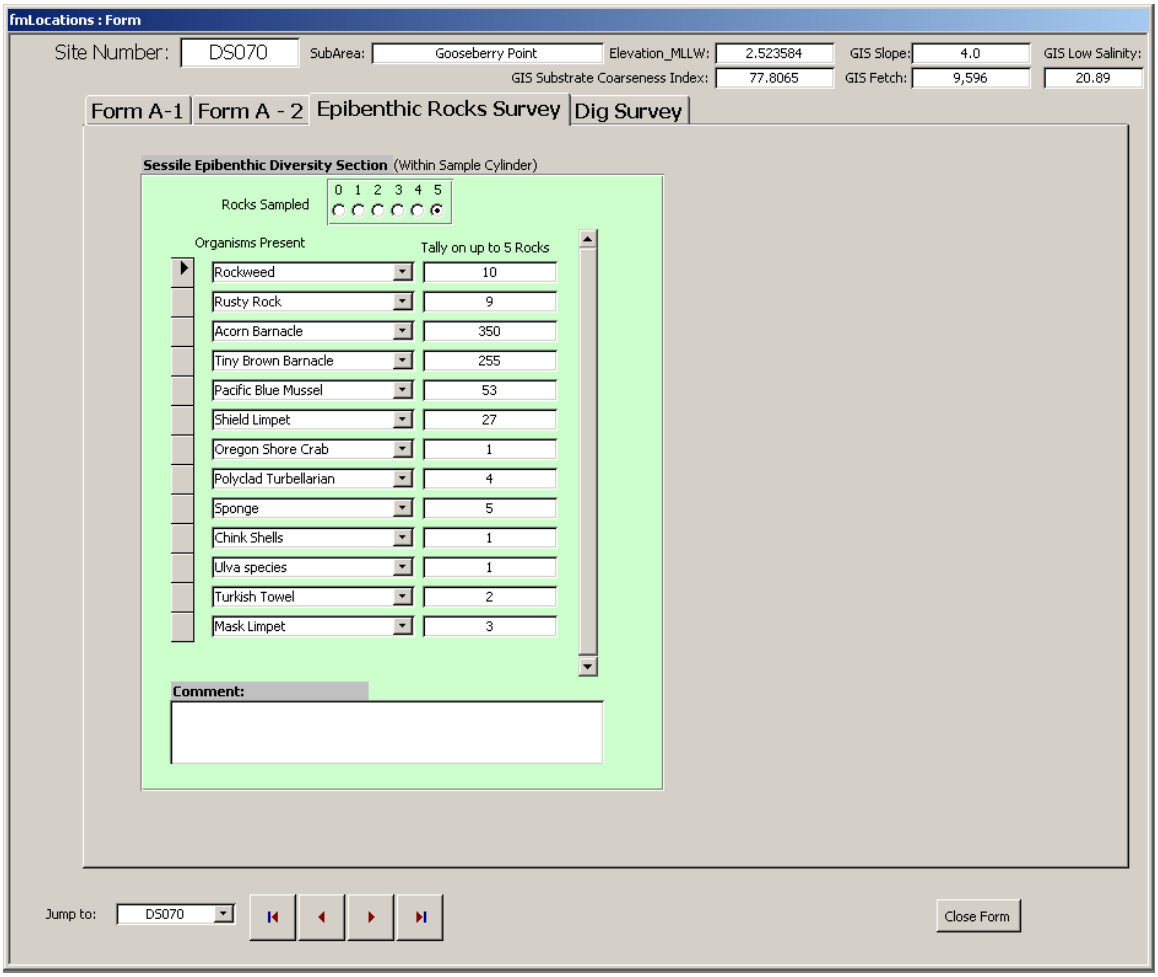

Figure J.8. Locations Form Showing the Third Results Tab for Location DS070

The third Locations Form tab (labeled *Epibenthic Rocks Survey*) lists the number of rocks that were inspected for epibenthic organisms, and the resulting list of taxa with their total abundance on those rocks (Figure J.8). The comment field visible on this tab relates specifically to the epibenthic rock survey at the site.

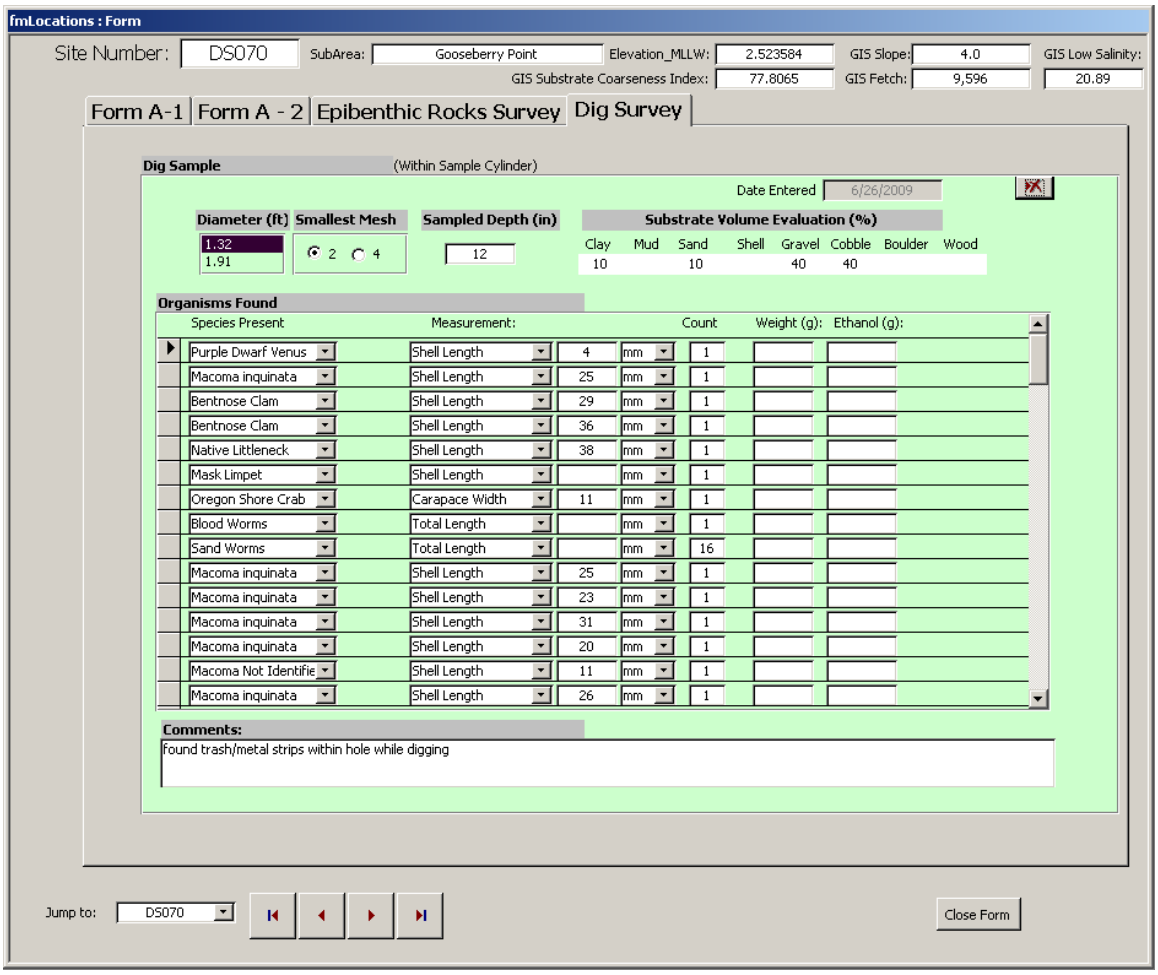

Figure J.9. Locations Form Showing the Fourth Results Tab for Location DS070

The final tab on the Locations Form (labeled *Dig Survey*) shows the results from the excavated sample taken from the site. The details of the sampled area include the size of the sampling cylinder used, the depth excavated, the smallest screening mesh used, and the subjectively estimated composition of the excavated substrate (Figure J.9). Below the sample information, the specific details of the organisms that were identified from the sample are also listed. The comments field visible on this tab relates specifically to the dig survey at this site.

To return to the main menu, click the *Close Form* button that is located at the bottom-right of the Locations form.

3.1.2.3 Exporting Summary Data to GIS

The button titled *Update GIS* exports summary results for each taxon based on the detailed information that is stored in the Dig Survey database. This summary data is exported into the table *DigSurveyGISOutput*, which is found in the database file *GIS_OutputTables.mdb*. The data in this table does not need to be updated again unless changes are made to the survey results in the database in the future.

The reason a separate access database was used to store the output data from this process is because the GIS software cannot access the data in the *DigSurveyDatabase* at the same time that the user is using the *DigSurveyDatabase*. Without an intermediary database, the GIS software would require that the user exit out of the *DigSurveyDatabase* before the data could be viewed.

3.1.2.4 Exporting Custom Data to GIS/Excel

The final button on the main menu is titled *Export Selection*. This button opens a form that allows the user to select sites based on geographic sub-area, elevation ranges, sampling month, or any combination of these criteria (Figure J.10). When the form opens, no sites will be listed on the form. In order to populate the list of sites, the user must first choose their site-selection criteria then click on the *Apply Filter Selection* button. If the user wishes to select all sites, the default filter settings should be left in place and the *Apply Filter Selection* button clicked.

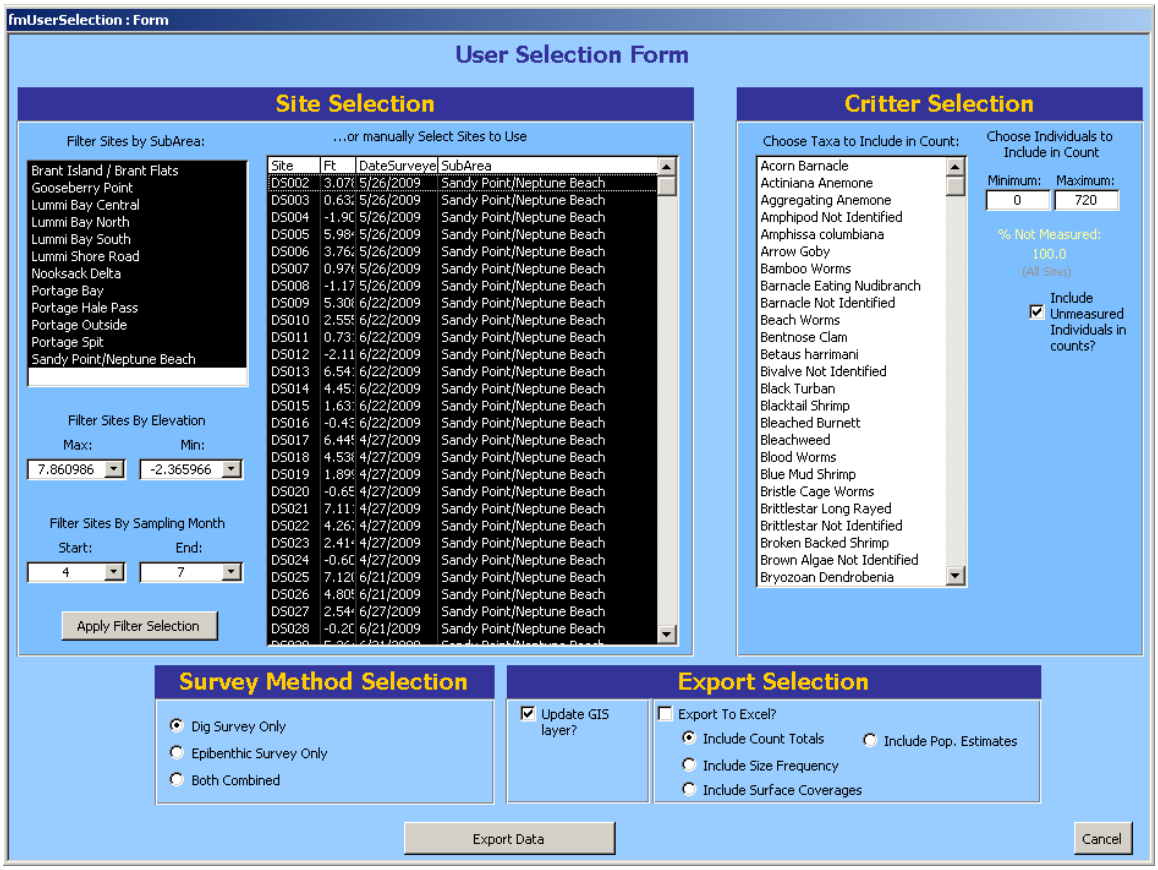

Figure J.10. Form Used to Export User-Selected Data From the Dig Survey Database

In addition, the user can select which taxa will be included in the output results. For example, the user may elect to combine Manila clams together with Pacific littleneck clams and obtain results for both taxa combined. To achieve this result the user would select both Manila clam and Pacific littleneck from the list of species (Figure J.11). If the results for only one species were desired, then the user would only select that single species from the list.

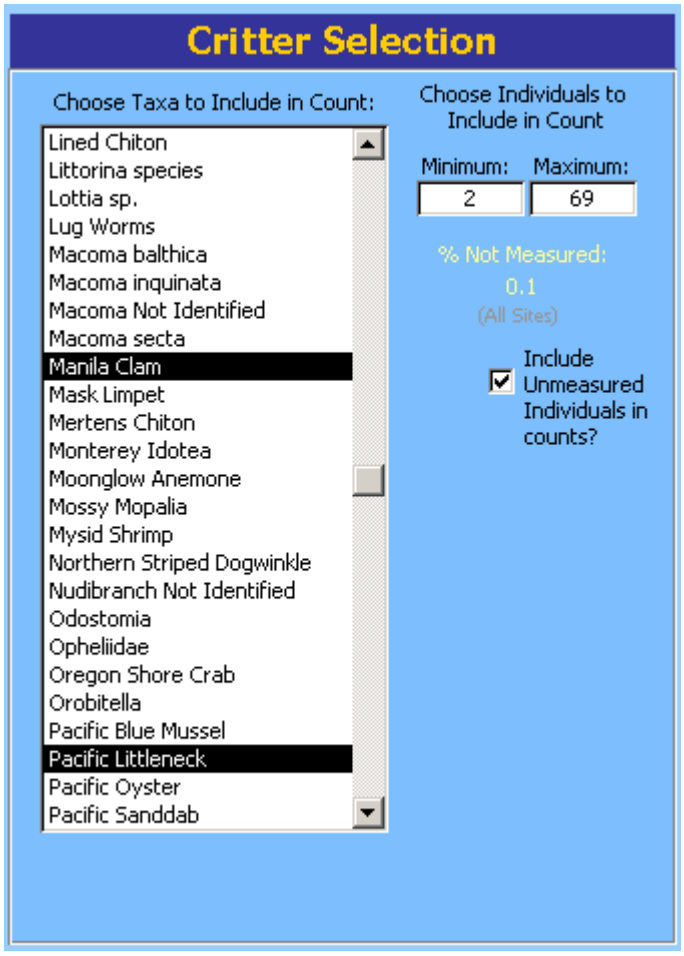

Figure J.11. To Combine Results For More Than One Taxon Select More Than One at a Time (Shown) To Limit the Results to Only One Species, Select Only One Species

The user may also select a size range so that individuals outside of that range are excluded from the results. This allows the user to explore questions about the distribution and abundance of juveniles separately from adults. However, because some individuals could not be measured, due to breakage or time constraints, the user is additionally required to choose whether to include or exclude unmeasured individuals from the counts. To assist the user with making that determination, the database calculates and displays the percentage of individuals from the selected taxa that were not measured.

Once the user has selected the sites to be exported and selected the taxa and sizes to be included in the results, the user can choose whether to combine the data from the dig survey with the epibenthic rock counts, or whether to export results from only one of these survey methods.

Finally, the user can select whether to export the data to the *Dig Survey GIS Project* (discussed in Appendix K) and/or to create a Microsoft Excel spreadsheet file containing the results. If the user elects to create a Microsoft Excel file, they can further specify which kinds of data they wish to have included in the file.

The *Include Count Totals* option creates a worksheet named *qryUserSelected_GISOutput* that contains a row of data for each site with multiple columns containing information about the sites such as latitude and longitude, date surveyed, the area excavated, and various environmental parameters that were recorded. In addition, there are six columns that describe the settings that were specified by the user at the time of export. The columns *CrittersIncluded*, *SizeRange*, and *UnmeasuredCritters* respectively contain the list of taxa included in the counts, the size range of individuals that were included, and whether unmeasured individuals were included or excluded from the count. The raw counts for the included individuals are present in the column *TotalCount*, and the resulting density (the count divided by the area excavated) is reported in the column *ObservedDensity*. If size-weight data were collected for the taxa, then the column *BiomassDensity* contains the estimated total biomass for the selected individuals divided by the excavated area.

If an individual specimen was weighed, this weight is used to represent the biomass of that individual. Based on comparisons of weight measurements taken on individual clams before and after sample preservation in denatured ethanol, a small correction factor was used to adjust the weight measurement for specimens that were only weighed after being preserved in ethanol. If a particular individual was not weighed, then the database estimates the biomass of that individual instead. Biomass estimates were calculated using the appropriate size-weight relationship for each species along with the measured size of each individual. If no size was recorded for an individual specimen, the database uses the estimated weight for the average size of all measured individuals from the same species. If no size-weight data were recorded for that species at all, the resulting biomass for that species was assigned a zero.

The option *Include Size Frequency* creates an additional worksheet that provides the data required to create a size-frequency histogram of the individual measurements for the selected taxa. There are two additional columns on this worksheet called *TotalObserved* and *TotalUnmeasured*. The first is simply a count of all individuals from the selected taxa that were encountered during the survey (regardless of whether they are within the exported size-range). The second lists the total number of individuals that were not measured.

The option *Include Surface Coverages* creates a third worksheet that lists the percentage of coverage for all surface coverage types encountered at the sites. Note that the percentages for each row may exceed 100% if summed. Substrate-type coverages (e.g., mud, sand, gravel, cobble) should sum approximately to 100% because these were assessed collectively. Other types of surface coverages were each assessed independently.

This is because algae, vascular plants, and epibenthic fauna might have been present as multiple overlapping layers, and each of these layers might have partially or completely covered the surface.

The last option, *Include Pop. Estimates*, creates two additional worksheets that are named *qryUserSelected_ThiessenStats_b* and *qryUserSelected_ThiessenStats_c*.

Both worksheets use the Thiessen polygon analysis method to calculate the total population of the selected taxa, along with the appropriate 95% confidence limits. The first worksheet estimates the total biomass for the selected size-range of the selected taxa. The second worksheet calculates the total abundance of individuals for the selected size range of the selected taxa.

When all export settings have been finalized, the user can click the *Export Data* button to start the export process. If the user elected to create a spreadsheet file they will be prompted to specify a file location and name during the export process.

3.2 Shorebird and Marine Mammal Database

3.2.1 Tables Summary

Data from the shorebird and marine mammal survey are located in the *ShorebirdMarineMammalsDB.mdb* database application. Figure J.12 shows the tables present in the file, and defines the relationships between the tables.

The tables of primary importance are *SurveyRuns*, *Locations*, *SurveyRunLocations*, and *SurveyRunLocationTaxonomicUnits*.

- Table *SurveyRuns* stores the date of each monthly sampling event, and table *Locations* stores information describing the locations and geographical sub area of each survey site.

- Table *SurveyRunLocations* stores information about the conditions during each site visit.

- Table *SurveyRunLocationTaxonomicUnits* lists the taxon id value and counts for taxa encountered during each site visit.

- The common names of each taxon are listed in table *Taxonomic Units*.

- The tables *TaxonomicGroups* and *TaxonomicCategories* contain allow different taxa to be grouped into relevant higher-level groupings.

- The remaining tables (*Personnel* and *SurveyRunPersonnel*) pertain to the personnel present during each sampling event.

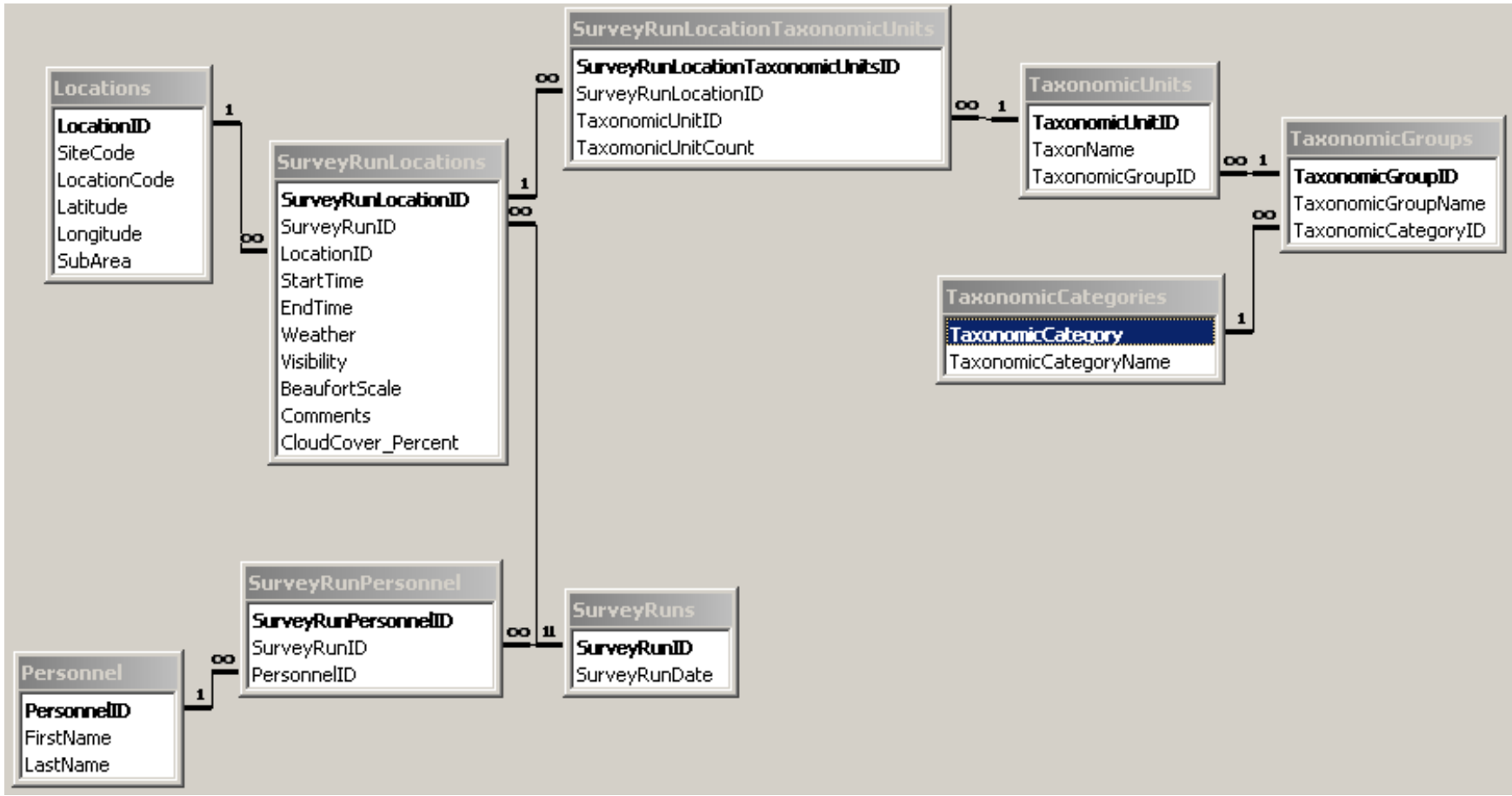

Figure J.12. Entity Relationship Diagram for ShorebirdMarineMammalsDB.mdb

3.2.2 User Interface

By default, the database window is hidden from view when the database opens. This is by design to prevent accidental erasure of data or database objects by inexperienced Access users. Experienced access users can view the tables and other database objects by un-hiding the database window from the Window menu.

Figure J.13 shows the main menu form that appears when the Shorebird and Marine Mammal Survey database is opened.

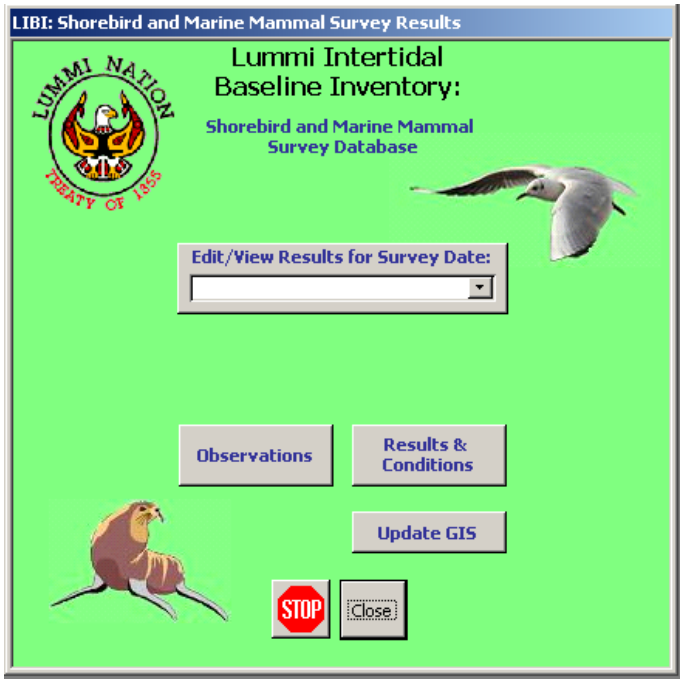

Figure J.13. Shorebird and Marine Mammal Survey Database Main Menu

The two buttons at the bottom of the menu will either close the database (*Stop*), or simply close the menu (*Close*) while leaving the database file open.

The two buttons immediately below the drop-down list each open a query view of the data that can be copied and pasted into other software such as spreadsheets.

The *Observations* query lists each taxon found at each site during the survey period in rows, and shows the counts for each sampling date in separate columns.

The *Results & Conditions* query arranges the data so that there is a row for each site visit, ordered by date, and separate columns for each taxon encountered during the survey.

The button titled *Update GIS* opens a form (Figure J.14) that allows the user to select one or more taxa to combine, and then exports the resulting counts to a GIS compatible format. The database exports the shorebird and marine mammal data into the table *BirdSurveyGISOutput* in a separate access database titled *GIS_OutputTables.mdb*. This data is used by the *Bird Survey GIS Project*, which is further described in Appendix K. To return to the main menu the user can click the *Done* button.

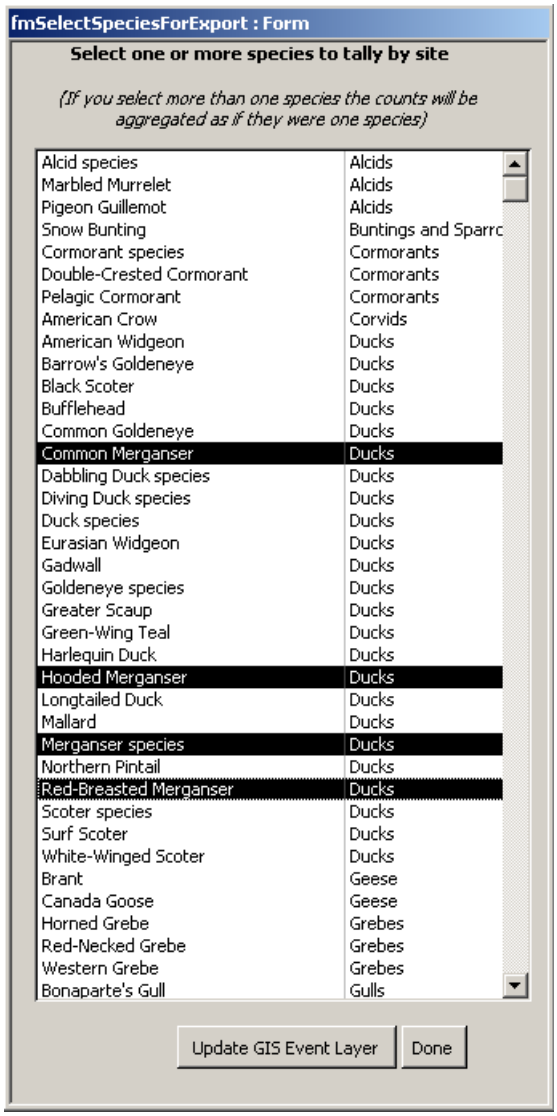

Figure J.14. The Update GIS Export Form is Used to Export Total Counts of the Selected Taxa to a GIS-Compatible Format.

On the main menu, the drop-down list titled *Edit/View Results for Survey Date* lists the dates when bird survey activities took place. When the user selects a date from the list, the survey run form opens and shows the survey results for the first site visit on that date (Figure J.15).

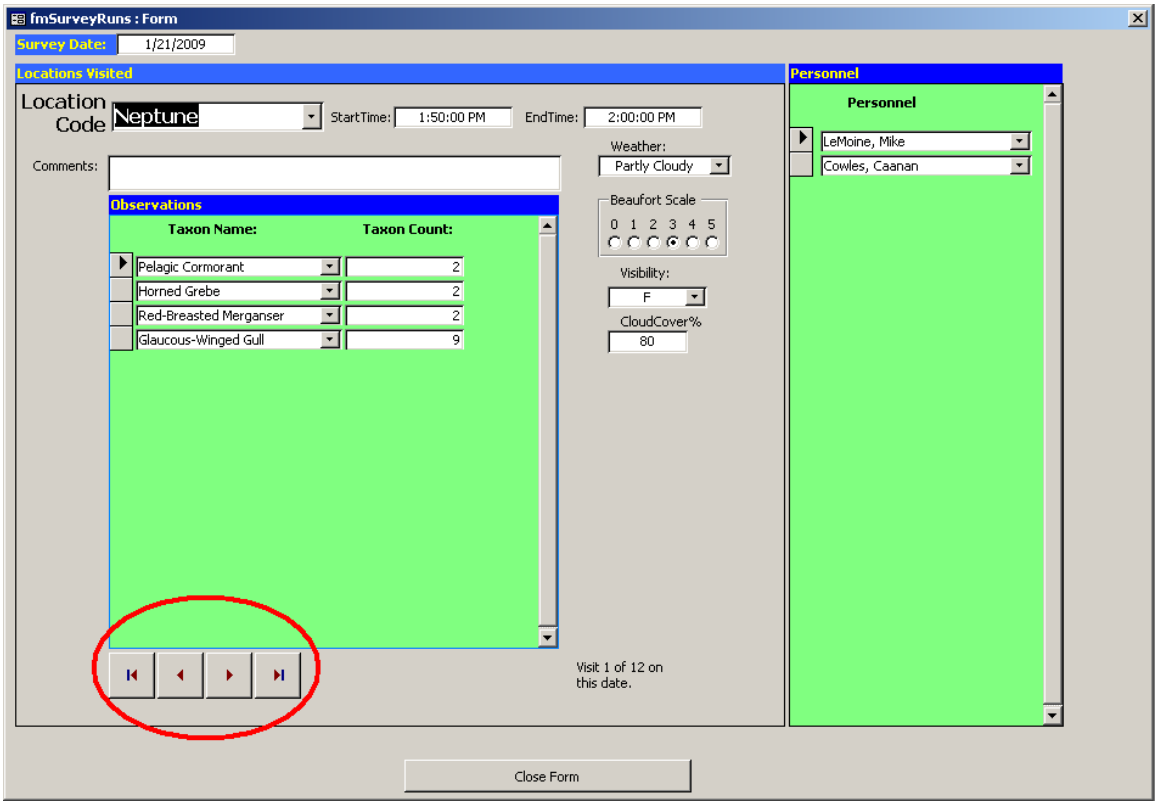

Figure J.15. Survey Run Form Showing Results from the First of 12 Site Visits on this Date (Navigation buttons are circled in red.)

Results for the remaining sites that were visited on that date can be viewed by clicking on the navigation buttons below the listing of taxa. To view data for another date, return to the main menu by clicking the *Close Form* button and select the desired date from the drop-down list.

3.3 Finfish Survey Database

3.3.1 Tables Summary

Data from the finfish survey are located in the FinFishDB.mdb database application. Figure J.16 shows the tables present in the FinFishDB.mdb file, and defines the relationships between the tables.

The tables of primary importance are *SurveyRuns*, *Locations*, *SurveyRunLocations*, and *Fishes*.

- Table *SurveyRuns* stores the date of each monthly sampling event. The related table *SurveyRunPersonnel* stores information pertaining to the specific personnel participating on a particular day.

- Table *Locations* stores information describing the location and habitat/geographical characteristics of each survey site.

- Table *SurveyRunLocations* stores information about the environmental and weather conditions that were present during each site visit. The related table *ProfileObservations* stores temperature and salinity measurements taken at different depths at each site visit.

- Table *Fishes* lists the taxonomic unit identification value and group count for each taxa encountered during each site visit. Metadata about the fishes encountered, such as a length measurement, estimate of gut fullness (%), or marks/tags present are also recorded in this table. Other types of data about individual fishes are stored in separate tables. For example, items found in the gut contents of fish are stored in the table *FishGutItem*, and the details of any samples taken for DNA analysis, coded wire tag readings, or otolith reading are stored in the table *FishSamples*.

- The common names for each taxon are listed in the table *Taxonomic Units*. Wherever a field name indicates that a code value was stored, the corresponding name is stored in the table *Codes*.

- The tables *TaxonomicGroups* and *TaxonomicCategories* allow different taxa to be grouped into relevant higher-level groupings.

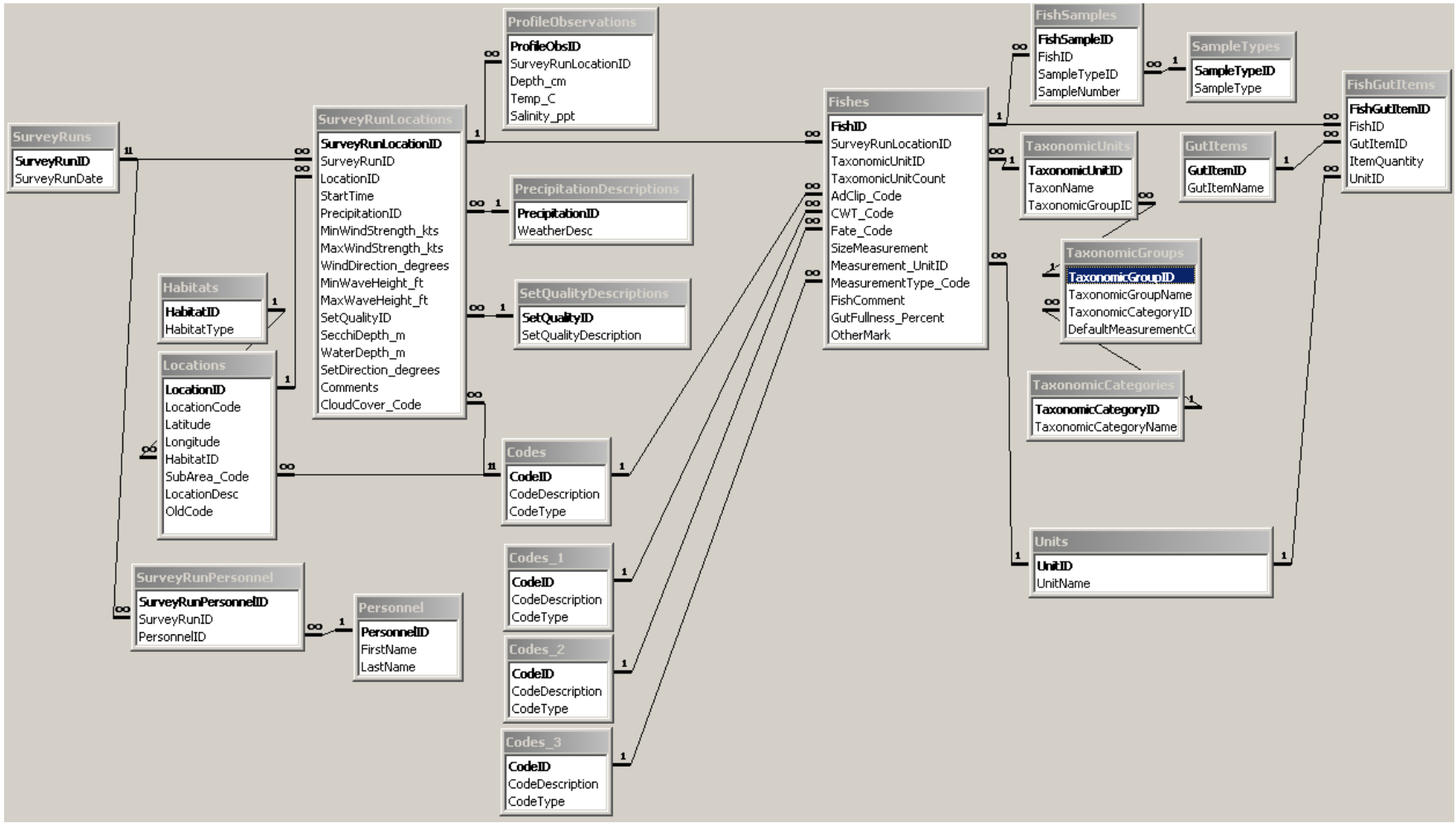

Figure J.16. Entity Relationship Diagram for FinFishDB.mdb

3.3.2 User Interface

By default, the database window is hidden from view when the database opens. This is to prevent accidental erasure of data or database objects by inexperienced Access users. Experienced access users can view the tables and other database objects by un-hiding the database window from the Window menu.

Figure J.17 shows the main menu form that appears when the Finfish Survey database is opened.

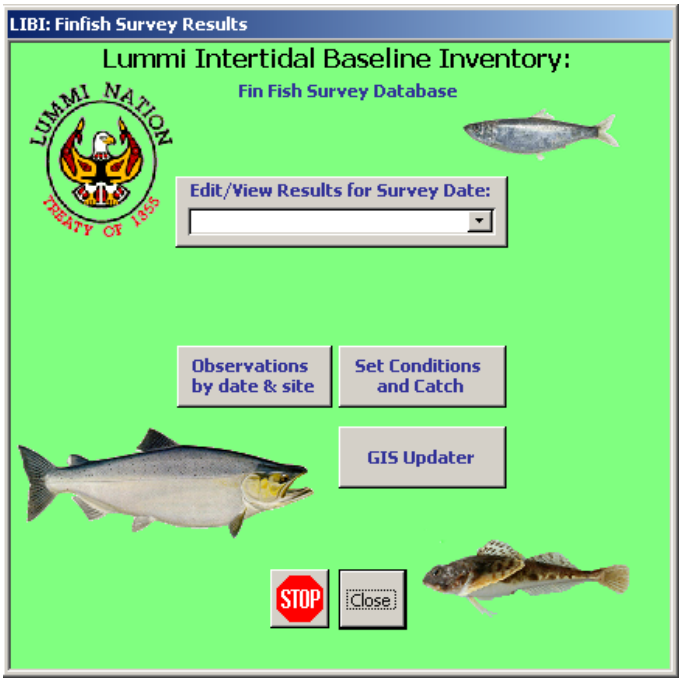

Figure J.17. Finfish Survey Database Main Menu.

The two buttons at the bottom of the menu will close the database altogether (*Stop*), or close the menu (*Close*) while leaving the database file open.

The two buttons that are immediately below the drop-down list each open a query view of the data that can be copied and pasted into other software programs, such as spreadsheets. If a site was visited on a particular date, the total catch for that taxon will be recorded as a number. If a site was not visited on that date, the result for that date will be null. The taxa encountered at each site are further subdivided into groups with various marking combinations so as to separate known hatchery-origin fish from presumptive wild-origin fish (no externally detectable marks).

The first of these queries, *Observations by date & site*, lists a row for each taxon that was encountered at each site, and shows the summary results in separate columns for each sampling date.

The second of these queries, *Set Conditions & Catch*, arranges the data so that there is a row for each site visit, ordered by date, and there are separate columns of counts for each taxon encountered during the survey. In addition, environmental conditions such as water depth, weather factors, and set quality of the Lampara net for each site visit are listed after the catch results.

The button titled *GIS updater* opens a form (Figure J.18) that allows the user to select one or more taxa to export summary results in a GIS compatible format.

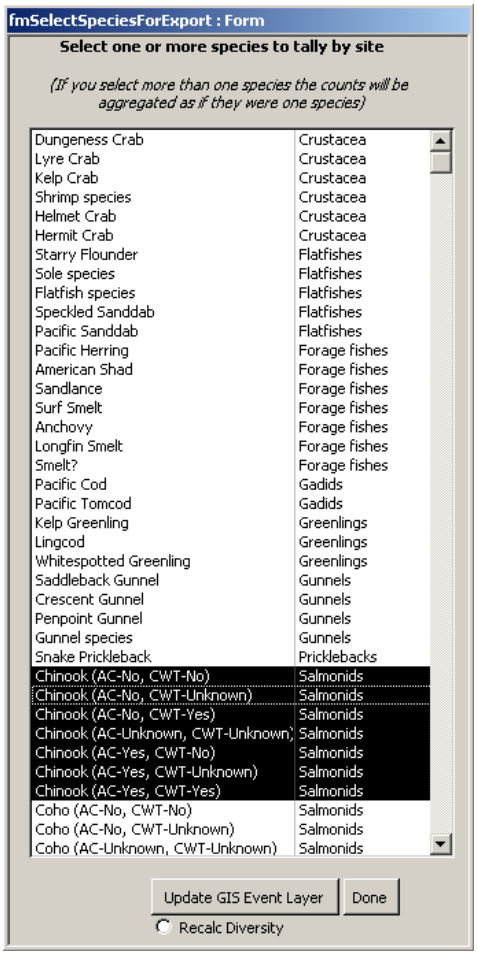

Figure J.18. The GIS Update Export Form is Used to Export Total Catches of the Selected Taxa to a GIS-Compatible Format

When the *Update GIS Event Layer* button is clicked, the database will export the selected summary results into the table *FinfishGISOutput* in a separate access database titled *GIS_OutputTables.mdb*. These data are used by the *Finfish Survey GIS Project*, which is further described in Appendix K.

The option to *Recalc Diversity* exports taxonomic diversity results to the table *FinfishGISOutput_TaxRich*, which is also located in the *GIS_OutputTables.mdb* database. Because this process takes several minutes to calculate, and the output already exists, it is not selected by default. This process should only be re-run if the survey data are modified.

The final item on the main menu is the drop down list of sampling dates. When the user selects a date from this list, the *Run Data* form opens to the first location that was sampled on that date (Figure J.19).

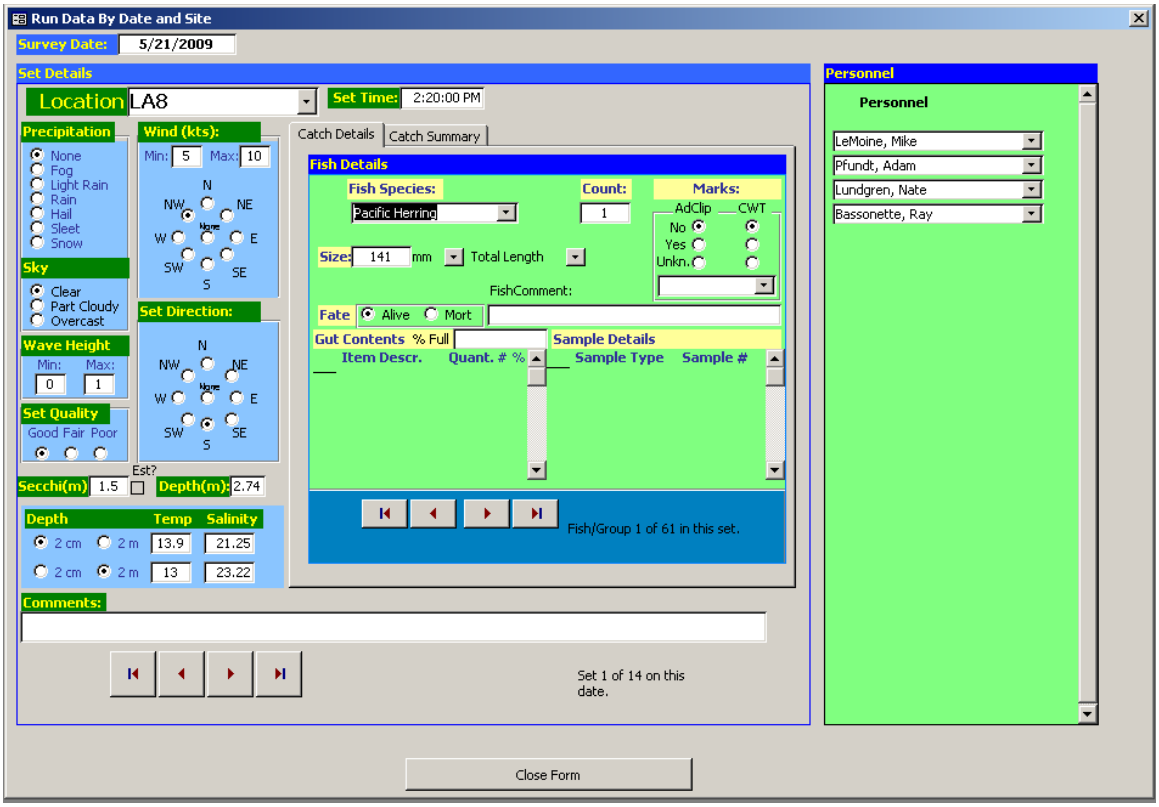

Figure J.19. Run Data Form, Showing the Detailed Results for the First Set Conducted on 5/21/2009

The Run Data Form shows information about the location of a set, the environmental conditions at the time that the set was conducted, and has two tabs that show the details and summary of the catch respectively.

The *Catch Details* tab shows the details associated with individual fish, or with groups of fishes. Details recorded for individual fishes may include the individual's size, mark-status, gut contents, samples taken, whether the fish was released alive or not, and any other fish-specific comments. The *Count* field details the number of individuals represented by each record. Users can view the details of other individuals or groups that were caught in the same set by using the navigation buttons supplied on the *Catch Details* tab.

The *Catch Summary* tab shows the total count for each taxon and mark-status caught during the set (Figure J.20).

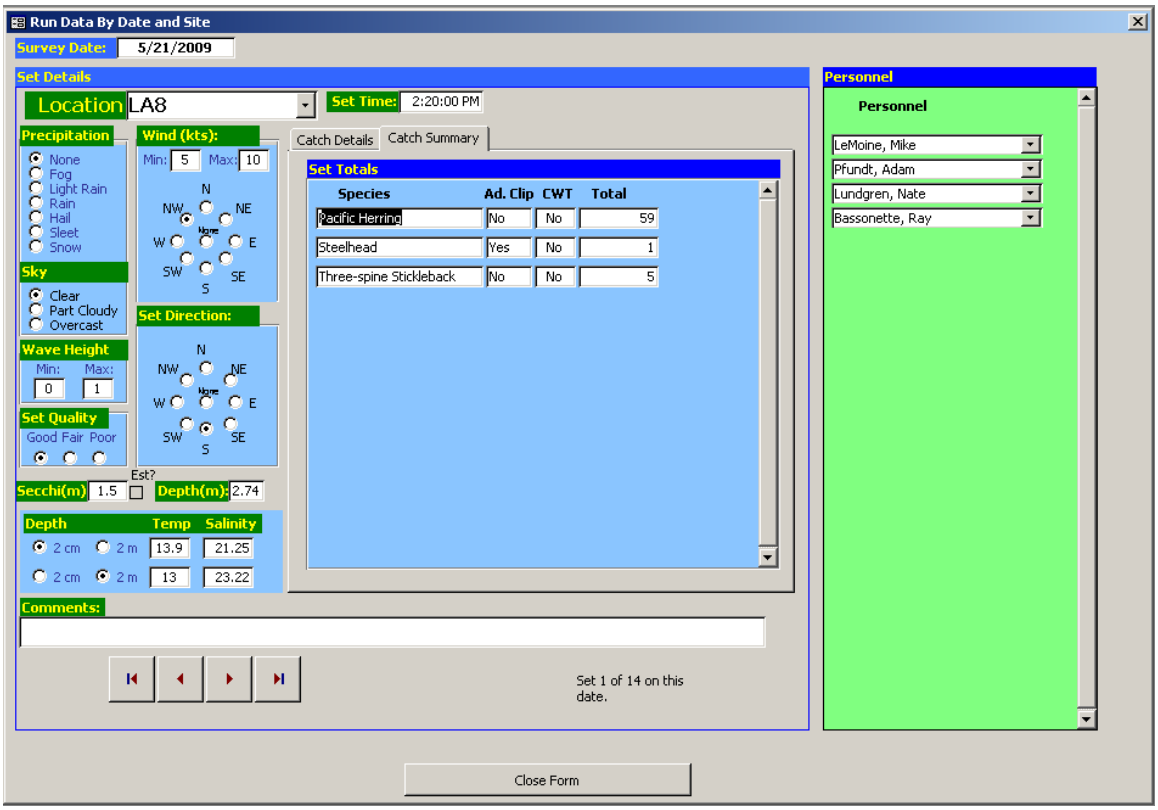

Figure J.20. Run Data Form, Showing the Summary Results for the First Set Conducted on 5/21/2009

To view results for a different location that was sampled on the same day, the user may click on the navigation buttons that are present below the *Comments* field.

To view results for a different date, the user can return to the main menu by clicking the *Close Form* button, and select a new date from the drop down list.

3.4 Large Bivalve Survey Database

3.4.1 Tables Summary

Data from the Large Bivalve survey are located in the HorseClamSurvey.mdb database application. Figure J.21 shows the tables present in the file, and defines the relationships between the tables.

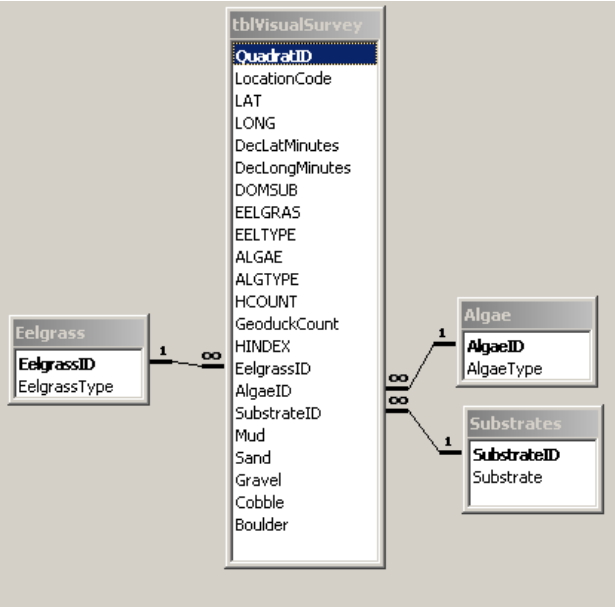

Figure J.21. Entity Relationship Diagram for HorseClamSurvey.mdb

Compared to the other database applications, the HorseClamSurvey.mdb is relatively simple. The most important table is tblVisualSurvey, which contains a row for each observation/count of horse clam densities in a 9-square foot (ft^2) area.

As described in Appendix B, most of these data points are simply counts of siphons that were present in a quadrat. However, many of the data points in Lummi Bay were derived using the transect method instead. The location of these points represents the midpoint of the individual sub-transects (Appendix B), and the value represents the average density that was observed in that sub-transect and adjusted to match the size of the quadrats used elsewhere.

The remaining tables exist solely to provide text labels to explain what the numerical codes stored in the primary table mean.

There are additional data points in the database that were not used in the final data analysis. These additional points fall into two categories. The first category includes a series of impromptu survey points that were collected on shore-parallel sand bars, which

were usually located near the subtidal fringe. However, these were excluded from the analysis because they greatly increased sampling density on relatively small features that were observed to have very high counts of horse clams. The second category of these additional points include a number of data points that resulted from a day of effort sampling horse clam densities on the fringes of Lummi Bay using an ad-hoc quadrat-based sampling strategy in an area of relatively high horse clam abundance. These data points were discarded because the sampling method was felt to be non-applicable over the larger extent of Lummi Bay, and that sampling methodology was discontinued in favor of using the transect methodology. These data are only included in the database for the sake of completeness.

3.4.2 User Interface

By default, the Large Bivalve Survey database window is hidden from view when the database opens. This is by design to prevent accidental erasure of data or database objects by inexperienced Access users. Experienced access users can view the tables and other database objects by un-hiding the database window from the Window menu.

Figure J.22 shows the data form that appears when the database is opened.

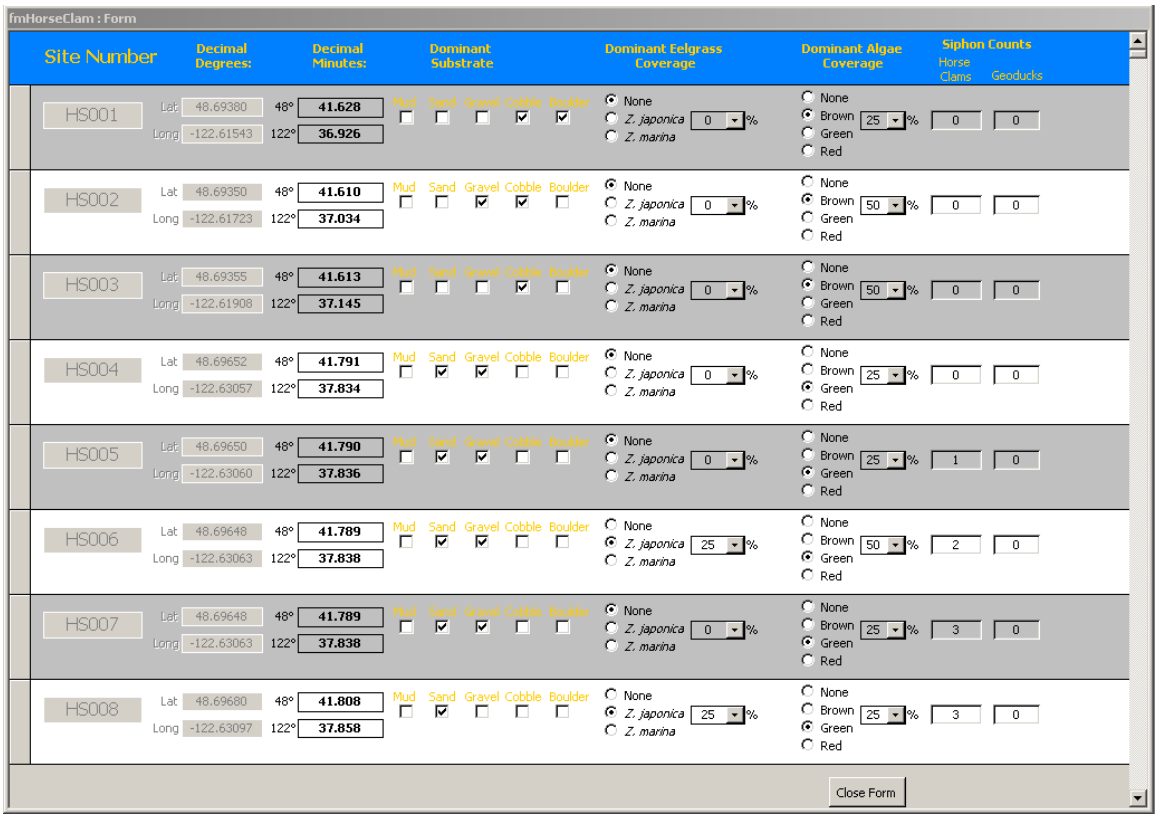

Figure J.22. Horse Clam Survey Data Form Showing Results for Each Observation of Horse Clam and Geoduck Clam Population Densities in a 9-ft^2 Area

The data for the Large Bivalve Survey required considerable post-processing in the GIS environment to separate out the topmost (biased) data points from the remaining data, and to exclude data from both the impromptu shore-parallel bar surveys and the ad-hoc sampling in Lummi Bay. Accordingly, there is no pre-built tool in the database to recreate the final data input to the analysis. However, the same make-table query (*qryOutput*) that was used to extract the input data used during the post-processing has been left intact. The two shapefile outputs that resulted from the GIS post-processing are located on the LIBI DVD in the folder \LIBI\GISPRojects\Shapefiles.

3.5 GIS Output Database

The GIS_OutputTables database was created to store data exports from the four main survey databases described above. The GIS software that was used to map the LIBI results has a requirement that no other software can access the data source at the same time that the GIS software is attempting to use the data. By locating the data source table in a separate database to circumvent that requirement, the user can have the primary database application open at the same time as the GIS program and use the database forms to modify the source data without having to exit the database or the mapping software each time.

There is no custom user interface constructed for this database because it exists solely to provide a tabular data source for use by the GIS software.

4.0 Discussion

The physical form of the results stored in these databases fills 9 three-ring binders and it was deemed to be impractical to attempt to include printouts of the raw data as attachments to this summary report. The original paper forms containing the field and lab data will be stored in the Archives Department of the Lummi Nation.

The Access 2000 database format used to develop these database applications is currently supported by Access 2000, Access 2003, and Access 2007 software. One of these programs is required to be installed on the user's computer to make use of the database applications. The database files are not protected from modification and future users may freely construct additional queries and forms in order to obtain data summaries that were not envisioned during the database development.

In the event that a post-spill survey is conducted that uses the same survey protocols, copies of the same database applications could be used to store and analyze the post-spill data with relatively little effort.

Storing the large quantity of raw data that were collected during the LIBI project in a relational database format, rather than a spreadsheet, was very helpful in reducing data management issues and enabled rapid data exploration and analysis for a very complex data set. It also allowed for interoperability with other analytical software programs such as ArcMap GIS and the R statistical platform that require tightly controlled data as inputs.

It is hoped that providing the raw data in an electronic format will make the data more accessible for further analysis and for comparisons with any future survey results.

This page intentionally left blank.