Desktop GIS for Environmental Sensitivity Mapping

Background:

In 1996, the first CD of digital environmental sensitivity index (ESI) data for Geographic Information Systems (GIS) was published by NOAA/HAZMAT. Geographic data were provided in ARC export format together with the data base files that had been utilized for data compilation and map generation. It soon became apparent that this format was only suitable for the high-end users, i.e. those who would be responsible for data base maintenance and updates. Significant manipulation and simplification of the data base was required by the users who wished to create and display custom map products, and those who wished to perform simple query and analysis of the data. This manipulation required users have access to sophisticated data base systems and that they become intimately familiar with a data base structure that is anything but intuitive. The consequence was duplication of effort by those who purchased the CD, erroneous interpretation of the data, and many who simply abandoned the idea of using ESI data in their GIS. The inability to use these valuable data was frustrating to both the users and NOAA. In an attempt to remedy the situation, HAZMAT has invested significant effort in developing a data set that would meet the needs of the growing number of desktop GIS users. This effort, which also involves additional quality checks for data consistency and usability, is discussed in the following pages.

Geographic Data:

Environmental Sensitivity data can be divided into three main categories: base map, biology and human use. The base map layers include ***hydro***, ***index***, and ***esi***. Biology includes some or all of the following layers: ***birds***, ***nests***, ***fish***, ***t_mammal***, ***m_mammal***, ***reptiles***, ***invert*** and ***habitats***. The human use layers are ***mgt*** and ***socecon***. For details on the content of these layers, see the Environmental Sensitivity Guidelines or the metadata for the atlas of interest. For diagrams of the data base structure reference this document or the metadata as the guidelines document does not reflect the current standard. All geographic data are provided in geographic coordinates with the horizontal datum being NAD 27. (An exception is Massachusetts which was published using NAD 83).

All geographic files are distributed in four different formats. The first of these is termed **SOURCE** data. These data are uncompressed, double precision ARC export files. These files most closely resemble the data as described in the Environmental Sensitivity Guidelines. Though they undergo the same quality control checks as all files, and in the case of the biology layers some minor modifications are made (see **Biology** heading), they are distributed in the same precision and hence have the resolution of the data used to generate the hard copy maps. Users who will be responsible for maintaining and updating ESI data should use these data in a GIS system that supports double precision data.

The second distribution format is found in the **AVPROJ** directory. This directory contains an ArcView project with a single view containing a theme for each ESI data layer. Themes have been symbolized using the standard ESI colors and patterns. It includes links from the biology and socioeconomic themes to a flat file version of the ESI tables. An ESI menu enables links to searchable breeding and source tables for more complex queries. These links can also be turned off to maximize the performance of simpler queries.

The third distribution format is **MOSS** formatted files. Moss is a simple ascii format that supports topology. Moss is supported by various mapping programs including the public domain program MOSS GIS and CAMRIS and its simple text format makes it easy to write translators to other GIS software. Users of these files should use the lookup tables provided in this directory (**BIO_LUT.TXT** and **SOC_LUT.TXT**) in place of the lookup tables in the DBFILES directory.

The final distribution format is found in the **ESI_VIEW** directory. Geographic data are provided in MARPLOT format bundled with the desktop data files in a standalone version of Filemaker. Installers are provided for both Macintosh and PC users to help load the esi application on your computer. TOUR should help familiarize users with the MARPLOT application as well as the content of the ESI data. It is hoped MARPLOT will meet the needs of users who wish to do simple query and analysis of the ESI data, but who do not have an established GIS/desktop mapping program in place. It may also be useful for other users to check data brought into their own mapping programs for proper linkage to the data base as well as general presentation considerations such as standardized color and hatch patterns.

Data Attribute Files:

In an attempt to support a wide range of users, data attribute files are distributed in both comma delineated text format and as uncompressed ARC export files. They are provided in a simplified "desktop" format, as well as a collection of data files for use in a relational data base as described in the metadata. With the exception of those responsible for updates to the ESI data, the simplified data format will meet the needs of most users. It should be recognized, however, that the desktop format is not normalized and changes made to the desktop files cannot revert to the relational structure. The desktop format is most suitable for users wanting to view, query and analyze the ESI data. On the ESI CDs, all data files are placed in a directory called DBFILES. In this directory there are two sub directories, desktop and relfiles each with two additional sub directories, text and export. The contents of the sub directories is identical information in varying formats.

Biology:
There are two primary desktop data files that link to the biology layers. The first, **BIO_LUT** should be used if you are working with a mapping program that requires each polygon have a unique id. This file provides a one to one link from the geographic layers. There are two versions of the **BIO_LUT** on each ESI CD the difference being the definition of the item *id*. Use the **BIO_LUT** file found in the DBFILE directory if you are working with the SOURCE or ArcView shape files. For the MOSS files, use the **BIO_LUT** found in the MOSS directory.

From **BIO_LUT** you will link to **BIOFILE**, or, if your mapping software supports many to many relationships from the geographic layers, you may link directly to **BIOFILE** via the *rarnum*. This file contains a summary of all the attribute data as well as links to two additional files **SOURCES** and **BREED_DT**. These additional files are provided for users who are using a data base that allows multiple links. The **SOURCES** file provides information about who collected the data, vintage of the data and publication information. This data is duplicated in the metadata file on the CD, though not in a polygon specific manner. The **BREED_DT** file provides life activity information in a searchable format. The utility of this file would be recognized when a user wants to select only the fish polygons where spawning is occurring in February. Though not in a searchable format, this breed activity is summarized in the **BIOFILE**.

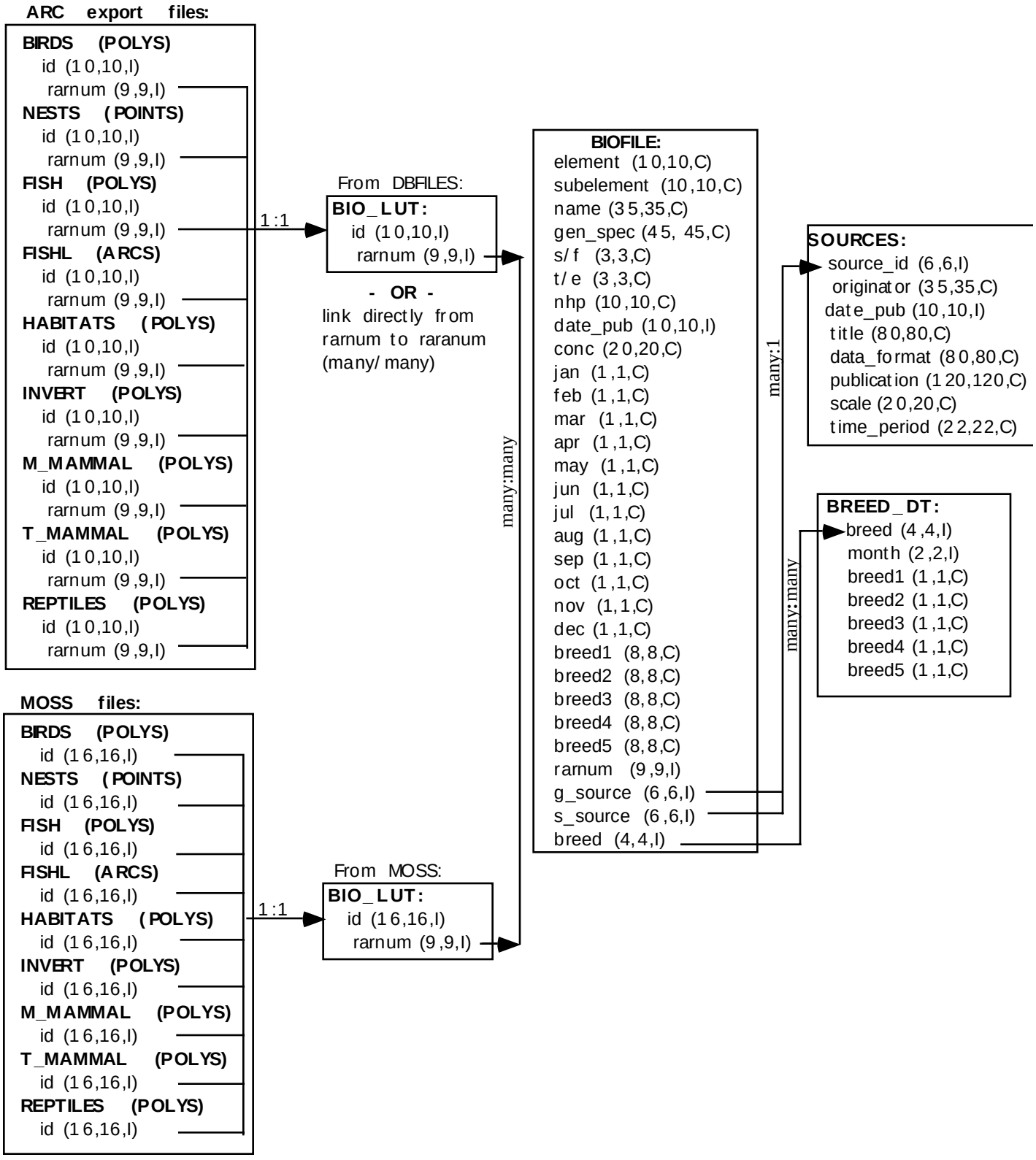

Figure 1: The relationship of the ARC and Moss files to the biofile

The *rarnum* in the source data files differs slightly from the *rarnum* described in the Environmental Sensitivity Guidelines. Digital files used to generate the hard copy maps are reprocessed to generate new rarnums that are specific to an element rather then passing through multiple layers. This supports the theory that digital files will be queried by layer. The element specific rarnum eliminates the confusion that would arise when reptiles appear as present in a polygon when a user is querying the bird layer. Once the new rars are generated, adjacent polygons that have the same species present (as well as the other criteria defining an rar) within a particular element, are joined to become one large polygon. Additionally, rarnums in the files distributed by NOAA will be unique across atlases. This allows the merging of multiple atlases and attribute files without the need to revert back to the always unique *id*. The *rarnum* is a nine digit number composed of the atlas number, layer/element number and the unique group number.

The first 19 items in **BIOFILE** (items element through dec) summarize information contained in the relational data tables **BIORES, SEASONAL, SPECIES, BREED** and **STATUS**. The items *breed1 - breed5* present a textual summary of the life activity data. This is presented for easy reference as well as to support the needs of users who are unable to link to the newly defined **BREED_DT** table provided. An explanation of each activity by element is described in the metadata. The last three items in the **BIOFILE** are the links to and from the **BIOFILE**. It should be noted that the rarnum key is a many to many link. There may be many polygons that have the same rarnum and there may be many records for each rarnum. At a minimum, there will be one **BIOFILE** record for each species present in an rarnum grouping.

Likewise the link to the **BREED_DT** table is many to many. Many species within the same element may share the same monthly life activities. Each breed variable will have twelve records in the **BREED_DT** table, one for each month, whether or not there is any special life activity present that month. Note that this is different from the **BREED** table in the relational data files where like activity schedules are duplicated for each species regardless of uniqueness and where records are not included if no special life activity is present for a month. The approach taken here disallows redundancy and thus reduces the size of the **BREED_DT** table significantly. It also assures all **BIOFILE** records link to a record in the **BREED_DT** table.

As previously mentioned, the **SOURCES** file provides information about who collected the data, vintage of the data and publication information. This table is exactly the same as the **SOURCES** table found in the relational data files. For additional information on the contents of the data fields, please reference the metatdata. There is a many to one link between the **BIOFILE** and the **SOURCES** file.

Figure 2 shows how to navigate through the relational files found in the relfiles directory. There is a link from the geographic layers to the data base either through *id* or the *rarnum* as well as the multiple links between the data files themselves. For the files **BIOFES, SOURCES, SPECIES, SEASONAL** and **STATUS** please reference the metadata for a description of content. Note that the lookup table here is **BIO_LUT**. This file combines all of the *id/rarnum* lookups regardless of whether they refer to point, polygon or arc data. The **BIO_LUT** file found in the relfiles directory is exactly the same as the **BIO_LUT** file used in the desktop version described above. Therefore, the rarnum is again unique across atlases as well as elements, as described above.

For more information about the relational tables **BREED** and **STATUS**, please see the metadata.

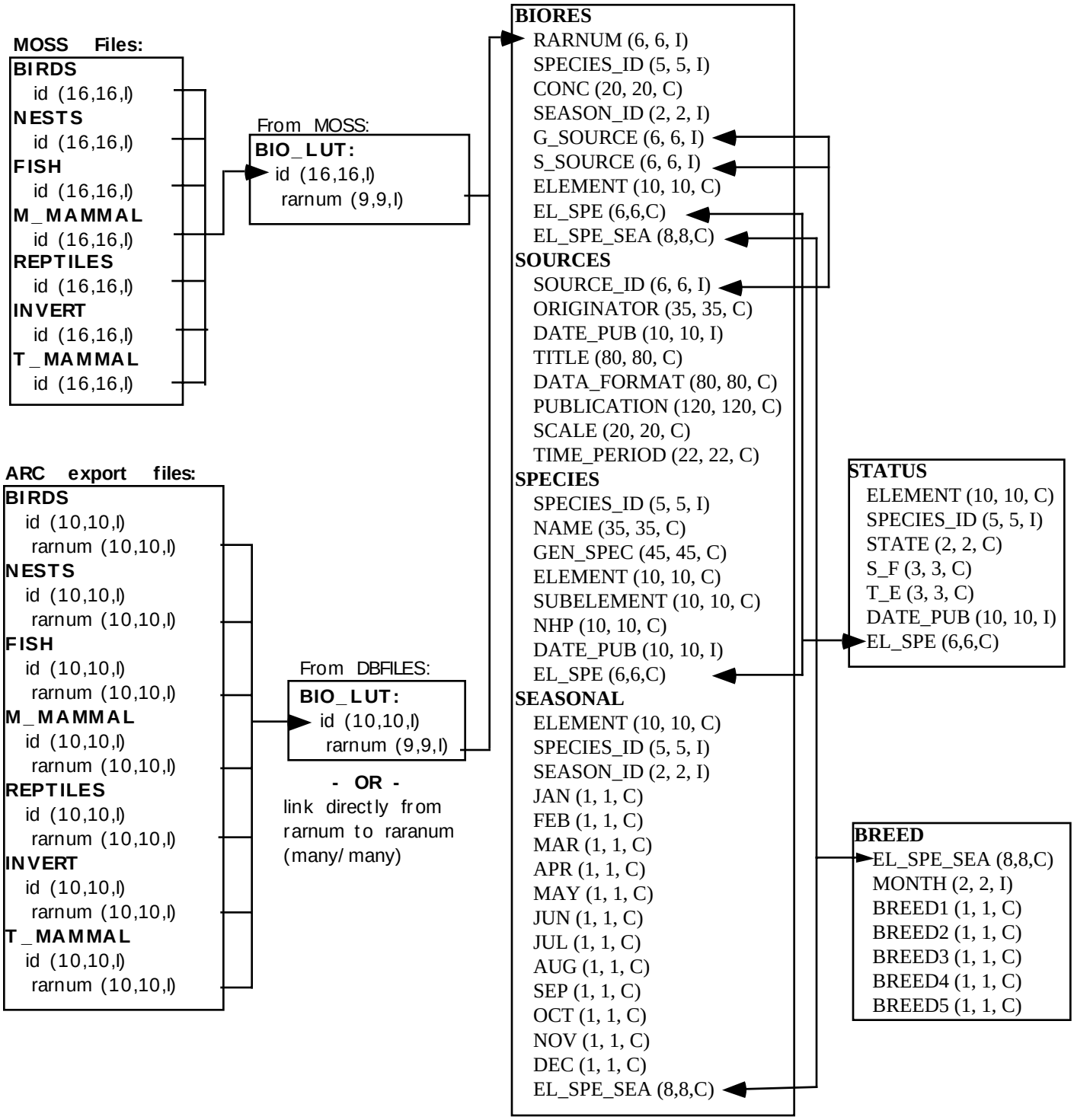

Figure 2: Navigating through the relational data structure

Human Use:
Attributes relating to the human use layers of the ESI are found in the file **SOC_DAT**. The link from the geography to this data table may be through either the *id* or *hunum* item. The *hunum* is the key id to a record that may describe multiple polygons or points found in the atlas. This number has been modified from that described in the Guidelines to be unique across atlases, similar to the type of modification discussed for the rarnum. Further information about what is contained in the **SOC_DAT** file can be found in the metadata.

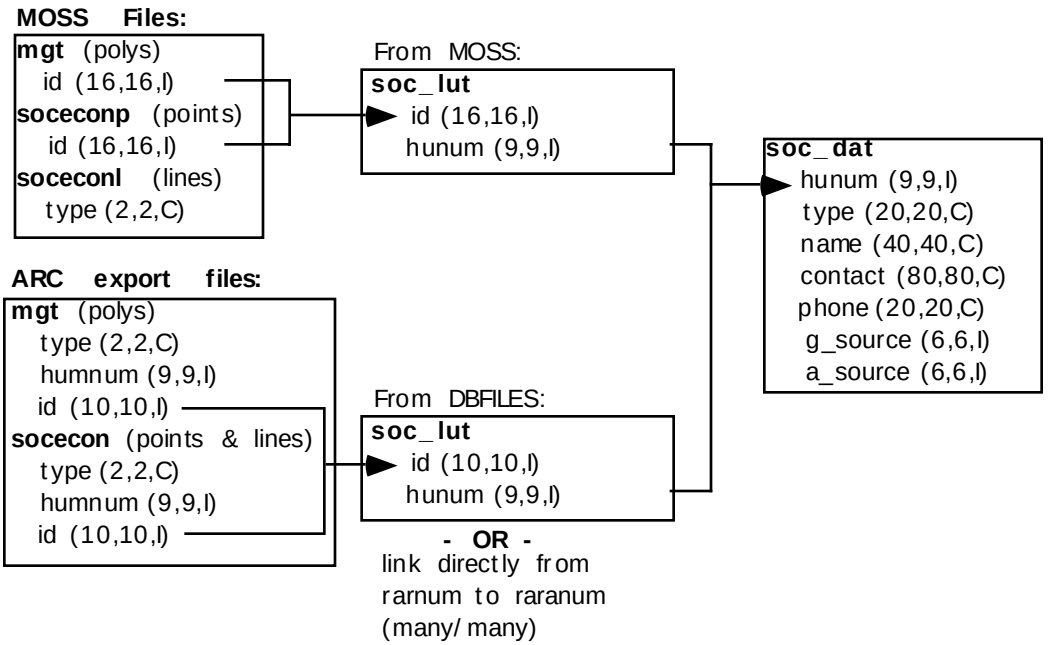

Figure 3: Linking to the Human Use attributes

Note in diagram 3 that the ARC files include an item called *type*. This field contains a two character abbreviation describing the kind of resource that's referenced. See the metadata for a list of abbreviations.

As mentioned earlier, MOSS files only allow the export of one attribute per file, so the item type is not included with the files that link to **SOC_LUT**. Another limitation of the MOSS file is it can contain only one feature type, i.e. points, lines or polygons. The ***socecon*** layer contains both points and arcs so each of these features is exported separately for the MOSS files. The line file, ***soceconl***, does not link to any other data tables so it is possible to retain the two character item "type".

Base Map Layers:
The base map layers, ***hydro***, ***esi*** and ***index*** do not link to any additional data tables, but there are certain items associated directly with the layers. The ARC export files retain all of the attribute data, however the MOSS files require selection of the highest priority attribute for export, as again, MOSS supports only one attribute per layer. Shown in figure 4 are the attributes associated with each of the base layers. The ARC files show the layers ***esi*** and ***hydro*** twice, once each for polys and once for lines. There will, however, be only one file for each of these layers. The poly attributes will be stored in the polygon attribute table (.pat) and the line attributes will be in the arc attribute table (.aat) The attributes for both the line and poly features will be available for use in ARC/INFO and possibly in some programs utilizing a translator. However, for ArcView 3 and MOSS it is possible to only

query/display one feature type per layer. Therefore, in order to access the esi line attribute values, it is necessary to export the cover with just a .aat file (***esil***) and just a .pat file (***esip***).

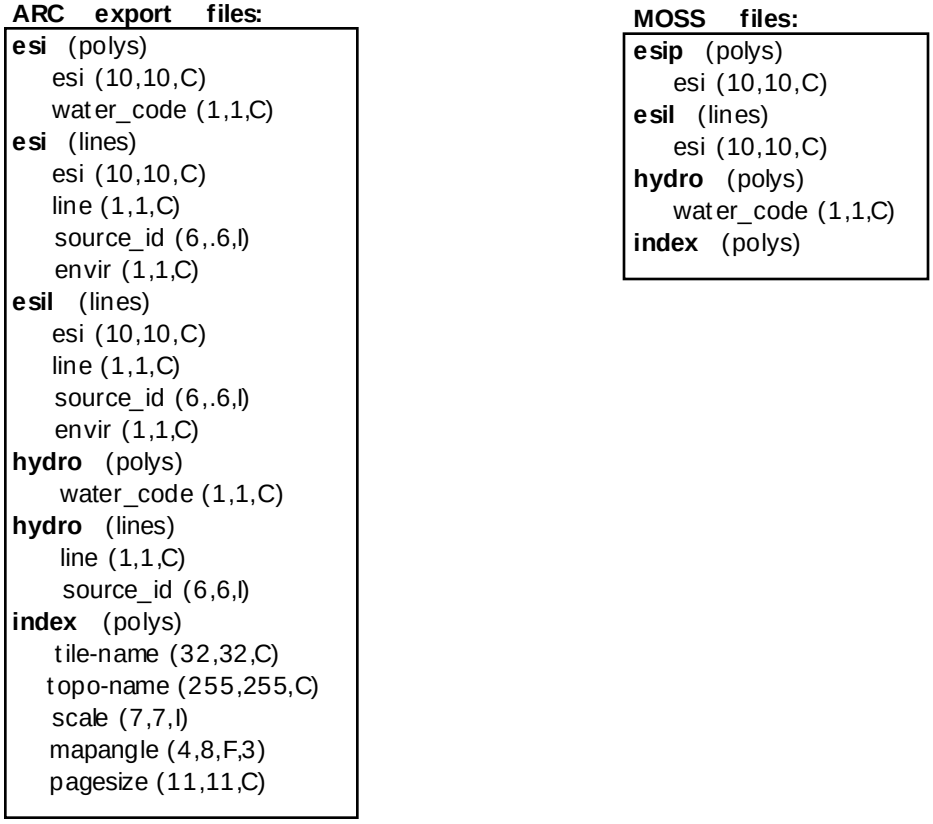

Figure 4: The base map layers and associated attributes

Figure 4 also shows which attributes were retained on export of the MOSS files. These values represent the crux of the information associated with each of the base layers For a description of the content of the base layer fields, please see the metadata. In the event an *esi* was undefined, the value for *water_code* was copied to the *esi* field. This means the water_code value in the MOSS files is basically retained since the water_code value is implicit through the esi value. Moving the *water_code* value to *esi* is also an aid when trying to display maps with the proper colors and symbology. See the next section for further details.

Displaying ESI Maps

One of NOAA's goals, when creating the ESI standards discussed in the Environmental Sensitivity Index Guidelines, was that users would be able to create their own maps using standardized colors and symbolization making them universally understandable by other ESI users. Extending this capability to the desktop is becoming more practical and some basic approaches are outlined below. Standard colors, hatch patterns and the ESI symbolset, are outlined in the Guidelines in chapter 5.

Base Map layers:
The ***hydro*** layer is color coded based on the value of *water_code*. Since the hard copy maps use the scanned quad sheets as a backdrop for the esi map, no standard color has been established for land or water. At this time it is suggested that the user select a subtle shade

of green to greenish-brown for the land and a shade of blue for the water. The values for *water_code* are 'W', water, and 'L', land.

Both the ***esi*** polygons and the ***esi*** lines are shaded based on the value of the item *esi*. RGB and CMYK colors are listed in chapter 5 of the Guidelines for all possible values of *esi*, with two exceptions. As explained in the previous section, *esi* may also be assigned a value of 'L' or 'W'. For these polygons, shade them using the same color scheme as was selected for the ***hydro*** layer.

Display difficulties arise when an arc is assigned multiple esi values, i.e. 10/3. For the hard copy maps, the width of the line is divided into the number of different esi values present, and, starting from the landward side, that portion of the line is shaded with the corresponding esi color. These maps are produced using ARC/INFO and an ARC symbolset and lookup table are provided on the desktop CDs for ARC users in the directory **SYMBOLS**. The solution in other mapping programs is less intuitive. ESRI expects that someday ArcView will be able to utilize ARC linesets and symbolsets which would solve the problem for those users. All users may want to contact the vendors of their particular software to see if they support custom line and symbolsets. Meanwhile, one "solution" is to code the the line with the color of the most sensitive esi value. The true value of the line could still be queried, but would not be immediately recognizable. In the ArcView project provided on the CD, a broken line means that shoreline section has multiple ESI values associated with it. The color reflects the most sensitive type present. In the ESI_VIEW project, shoreline segments with multiple ESI values are color coded with all of the ESI colors present, but not necessarily defined in the landward to seaward direction. To see the order of the ESI types present, the user can select the segment & check the value displayed in the bottom left corner. This display will be ordered from landward to seaward.

The ***index*** layer basically shows the boundaries of the USGS quads that were used when the hard copy maps were produced. Displaying the ***index*** simply as unfilled polygons may provide a locational reference to users. It also will help those who want to cross reference their digital data with the original atlas.

Biology layers:
The color and hatch schemes for the various biological layers are also outlined in chapter 5 of the Guidelines. Basically there is one color and one hatch pattern per layer. On the hard copy maps, if polygons of multiple elements overlapped, the polygon would be shown as black. If a user wants to print a map, they may want to try and emulate this practice. However for those wishing to use the digital data for query etc., the overlapping colors will be of little consequence.

It's important to realize that while a version of these digital data were used to produce the hard copy ESI maps, it was not without significant manipulation to the digital data causing them to appear differently and, in some cases, resulting in loss of information. For example, a hard copy ESI map usually has at least one common throughout box. For each of the rarnums represented in that box, a polygon has been removed from the map itself. This results in less clutter and easier to read maps. However, if someone wanted to query a GIS system to find out the areal extent of endangered turtles, the answer wouldn't include the polygon that was moved to the common throughout designation. It helps to remember the analysis and query functions are a very significant part of desktop GIS.

A user may want to make a customized map, for example, showing the polygons where the endangered turtles may be found in February. Once the polygons are displayed in their GIS program, they may want to go in and edit the data to create a more appealing final product. This may involve creating common throughout boxes, dissolving adjacent polygons and/or

adding symbolization. The ability to perform edits will be closely tied to the mapping software being used.

An ESI symbolset is provided on the desktop CD. For ARC users included is the entire symbol, line and shadeset files, as well as a lookup table that will help in color coding the ESI layer. A readme file is included for those users. There is a subset of the ESI symbols available in ArcView 3. A print out of the symbolset is shown in chapter 5 of the Guidelines.

Human Use:
The human use symbols should reflect the value listed in type. Reference appendix B of the Guidelines to see the valid values for socecon and mgt item type. These can then be assigned the appropriate symbols from the esi symbolset shown in chapter 5.

Reference:
NOAA. 1997. Environmental Sensitivity Index Guidelines, Version 2. NOAA Technical Memorandum NOS ORCA 115. Seattle: Hazardous Materials Response and Assessment Division, National Oceanic and Atmospheric Administration.