Testing the ArcGIS Marine Data Model: Using Spatial Information to Examine
Habitat Utilization Patterns of Reef Fish along the West Coast of Hawaii

by
Alyssa A. Aaby

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Presented June 8, 2004
Commencement June 2005

Committee in charge:
Dawn J. Wright, Major Professor (Geosciences)
Laurence Becker, Minor Professor (Geosciences)
Brian N. Tissot, Committee Member (Env. Science, WSU-Vancouver)
Arthur W. Sleight, Graduate Council Representative (Chemistry)

AN ABSTRACT OF THE THESIS OF

Alyssa A. Aaby for the degree of Master of Science in Geography presented on June 8, 2004.
Title: Testing the ArcGIS Marine Data Model: Using Spatial Information to Examine Habitat Utilization Patterns of Reef Fish along the West Coast of Hawaii.

Abstract Approved:

_____

Dawn J. Wright

In response to anthropogenic pressures that have degraded habitat and put marine resources at risk (Leslie et al., 2003; Mumby et al., 2001; Puniwai et al., 2003), there has been a growing interest in the use of marine protected areas (MPAs) as a management tool to help slow, prevent or reverse negative anthropogenic changes. Recent studies have shown that the most effective MPAs are those which represent a full range of habitat types (Leslie et al., 2003; Carr et al., 2003). Yet, available scientific research has not evaluated the near-shore marine habitat utilization along the West Coast of Hawaii at the large-scales utilized by resource managers. Thus, this study focused on identifying the regional habitat utilization patterns for selected Hawaiian reef fish species to determine the most effective combination of habitat types. In addition, the habitat utilization analysis was used as one of several case studies to test the ArcGIS Marine Data Model's (MDM) (Wright et al., 2001) adaptability to work with real-world data and perform real-world analyses, as well as meet the five goals outlined by the MDM Working Group (Wright et al., 2001).

Created in 2001, by researchers from Oregon State University, Duke University, NOAA, the Danish Hydrologic Institute and ESRI, the MDM is a geodatabase template tailored to meet the needs of the marine GIS user community. Ultimately, this analysis will aid marine managers as it establishes correlations between small and large-scale habitat information, which provides a regional look at habitat utilization. Also, by testing the functionality of the MDM, its strengths and weaknesses will be identified so that it can be improved to better serve the marine GIS user community.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF TABLES

Master of Science thesis of <u>Alyssa A. Aaby</u> presented on <u>June 8, 2004</u>.

APPROVED:

_____

Dawn J. Wright, Representing Geography

_____

Chair of the Department of Geosciences

_____

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis of any reader upon request.

_____

Alyssa A. Aaby, Author

Testing the ArcGIS Marine Data Model: Using Spatial Information to Examine Habitat Utilization Patterns of Reef Fish along the West Coast of Hawaii

Introduction

Increasingly, anthropogenic pressures such as over-fishing, coastal development and aquaculture have degraded habitat and put marine resources at risk (Leslie et al., 2003; Mumby et al., 2001; Puniwai et al., 2003). In response, there has been a growing interest by resource managers, policy makers and academics in the use of marine protected areas (MPAs) as a management tool to help slow, prevent or reverse negative anthropogenic changes. Currently, more than 100 MPAs have been established worldwide; however, these protected areas encompass less than 1% of the world's oceans, less than 1% of U.S. waters (PISCO, 2002) and less than 1% of the main Hawaiian Islands (Tissot et al., 2004). The current MPAs in Hawaii cover approximately 1,236 km^2 (NOAA, 1999a), and consist of several different levels of protection such as the Marine Life Conservation Districts (MLCDs), Fishery Management Areas (FMAs) and Fishery Replenishment Areas (FRAs) (see Figure 1).
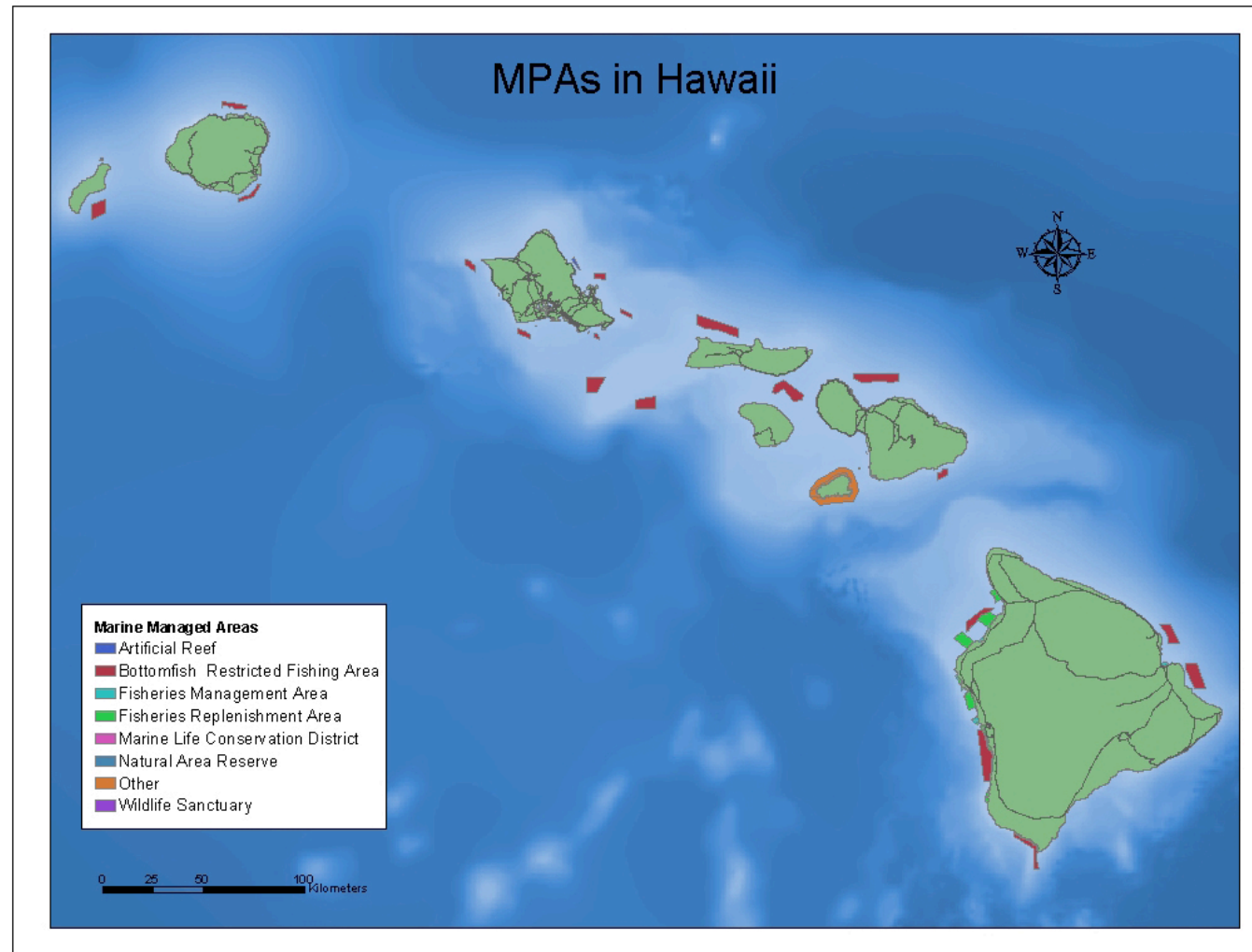
Figure 1. Map of MPAs around the main Hawaii Islands. The data used in this map are from the Hawaii Stateside GIS Program (DBEDT, 2004), NOAA (NOAA, 1999a) and WHAP (Tissot et al., 2004).

The most recent scientific research suggests that MPAs considerably enhance the conservation of marine biodiversity and contribute to the management of marine fisheries (e.g., Carr et al., 2003). These studies show that MPAs meet the economic and social needs of fisheries and communities dependent on ocean resources by protecting spawning grounds and increasing the abundance of fish stock in adjacent areas due to spill-over effects (e.g., Christensen, 2003). MPAs also preserve ecosystem components, such as habitat, food and shelter, which are critical to fish growth and survival. For example, it was shown that regardless of habitat quality, most protected areas in Hawaii had higher assemblage characteristics than open access areas and areas with less restrictive management regimes (Friedlander and Brown, 2003). Furthermore, MPAs in Hawaii have been demonstrated to effectively promote the recovery of fish stocks depleted by fishing pressures, without significant declines outside of reserves (Tissot et al., 2004). Thus, MPAs offer potential management solutions for resource protection and the prevention of over-exploitation.

Recent studies have also shown that the most effective marine reserves are those which represent a full range of habitat types and species assemblages, as well as take into consideration economic and social needs (Leslie et al., 2003, Carr et al., 2003). This study will focus on identifying the regional habitat utilization patterns for selected Hawaiian reef fish species on the west coast of the island of Hawaii (hereafter, West Hawaii). Specifically, this study utilized three spatial data sets: (1) The West Hawaii Aquarium Project (WHAP) fish survey data set (Tissot et al., 2004), (2) The National Oceanographic and Atmospheric Administrators (NOAA) large-scale habitat delineations (NOAA, 1999) and (3) The WHAP small-scale substrate data set,

to examine the regional habitat utilization patterns of 13 reef fish in the Reef/Aggregate Coral (RA) and Reef/Colonized Volcanic Rock/Boulder (RCB) habitat types (NOAA, 1999). To answer the question of habitat utilization, three sub-questions were investigated: (1) How do specific fish species utilize the RA and RCB habitats? (2) Does depth play a significant role in the classification of RA and RCB habitat types? And (3) Is small-scale substrate information obtained from WHAP correlated with large-scale NOAA habitat information?

The study of habitat utilization patterns is important because of the relationship that exists between fishes and their habitats (Christensen, 2003). Thus, by determining which combinations of habitat types are necessary for survival, the efficacy of the network of MPAs in West Hawaii can be evaluated. Additionally, and perhaps more importantly, available scientific research has not evaluated the current status of near-shore marine habitat utilization in West Hawaii at the large scales utilized by resource managers. Typically, most marine studies have been conducted at very small-scales; however, management units are usually on the scale of an island or an entire state and resource evaluation should reflect this scale (Friedlander and Brown, 2003). In order to address the knowledge gap that exists in habitat utilization patterns of common reef fish at a large-scale, the ArcGIS Marine Data Model (MDM) (Wright et al., 2001), a geodatabase template, was used.

Over the last three years, the GIS community has increasingly been engaged in building application-specific data models for ArcGIS in collaboration with the Environmental Systems Research Institute (ESRI) (ESRI, 2004). Some examples of

past application-specific data models include ArcHydro, Transportation

(UNETRANS) and Parcels (ESRI, 2004). These models were created in response to

the need for a better understanding of ESRI's new geodatabase data structure, and

better spatial modeling capabilities that capture the behavior of real-world objects

using object-orientation. In addition, the application-specific models allow for easier

data storage, assembly, and data sharing, while also focusing the model to meet the

needs of a specific community of users. Consequently, in 2001, researchers from

Oregon State University, Duke University, NOAA, the Danish Hydrologic Institute

and ESRI began work on a data model tailored specifically for the marine community.

The result was the first major draft of the MDM, which provides a geodatabase

template for marine users (Figure 2).
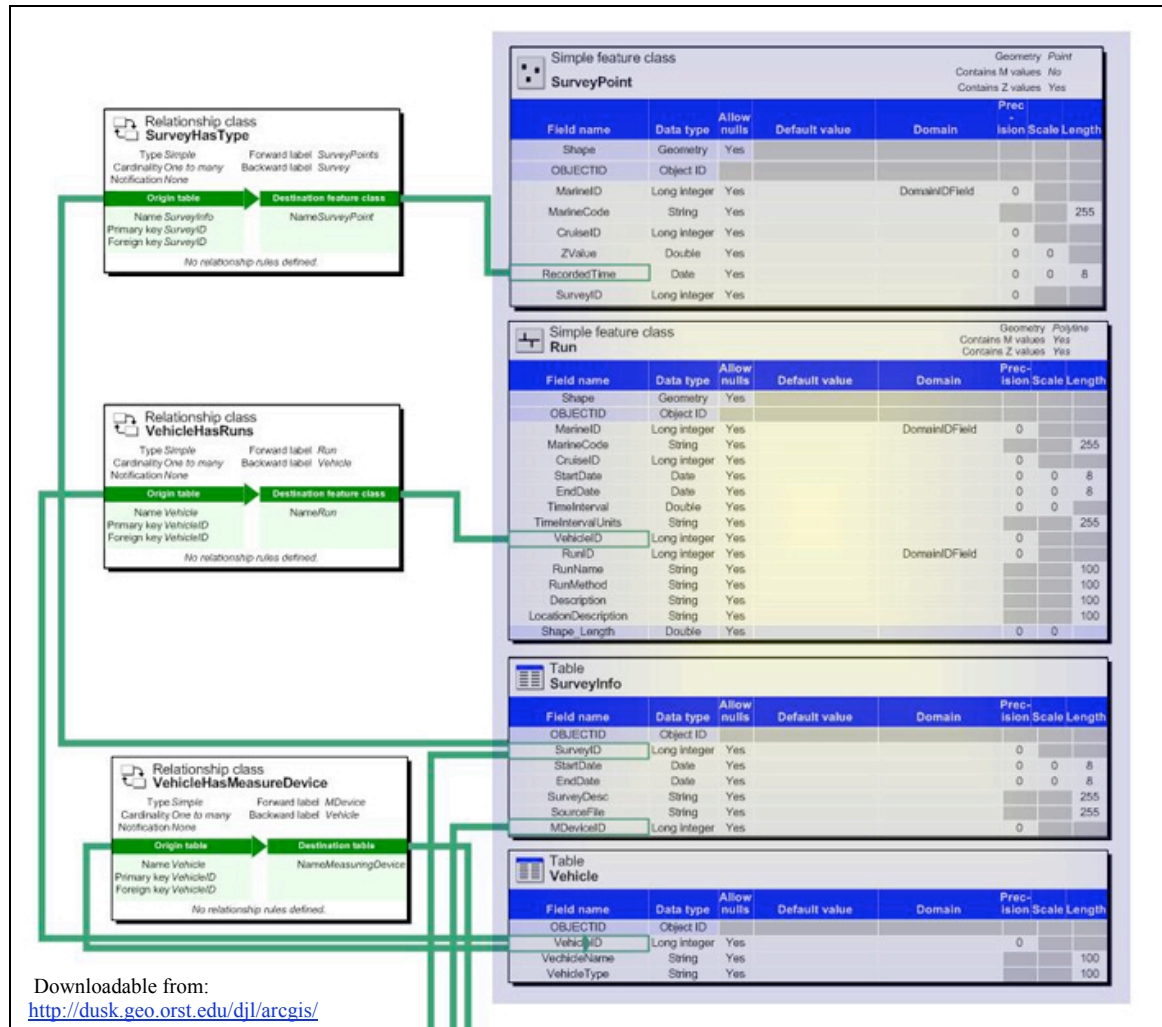
Figure 2. A portion of the geodatabase structure of the MDM. This shows a portion of the feature classes, relationships and tables in the model.

        The ArcGIS MDM was designed to be used as a geodatabase template for marine GIS users. The geodatabase template, like all geodatabases, is an organized hierarchy of data objects. These data objects are a collection of feature data sets, feature classes, object classes and relationship classes (Figure 3).

A geodatabase is the top level unit of geographic data. Points, lines, polygons can be housed in a geodatabase, when located within the spatial extent of the geodatabase.

A feature dataset. A feature dataset is a collection of feature classes that share a common coordinate system.

A feature class. A feature class is a collection of features with the same type of geometry (ie, point, line or polygon).

Metadata will be created both automatically and manually for any data entered into each feature class.

An object class. The onject classes store non-spatial information.

A relationship class is a table that stores relationships between features or objects in two feature classes or tables.

Figure 3. The organizational schema of the ESRI geodatabase.

Specifically, a feature data set is a collection of feature classes that share a common spatial reference. The spatial reference is part of the definition of the geometry field in the database. Accordingly, a set of transect survey points stored in the coordinate system NAD84 UTM Zone 4 could not be in the same feature data set as geographic latitude/longitude coordinates. In the geodatabase, all *objects* represent a real world object such as a marker buoy or lighthouse, and are stored in a row in a relational database table. Object classes are not represented geographically; however, they can be related to spatial information through a relationship class. Conversely, all of the *features* in a geodatabase are geographic objects that have a defined spatial location.

Basically, a feature is just like an object but it also has a geometry or shape column in its relational database table. The hierarchical data structure of the geodatabase allows the feature classes to inherit all of the attributes and behaviors of the object, but retain the spatial capabilities (Zeiler, 1999). The MDM geodatabase can store a range of data sets, from the small to medium data sets of personal geodatabases, to the very large geodatabases managed with the help of ArcSDE (Arc Spatial Database Engine).

To ensure that the MDM meets the needs of the marine community, the habitat utilization analysis was used as one of several case studies to test the model in its adaptability in working with real-world data and performing real-world analyses, as well as identify its strengths and weaknesses in meeting the five major goals. In particular, this habitat utilization analysis will test the ability of the model to meet goals one and two. However, the remaining three goals were also tested because their success is important for future habitat utilization studies, as well as the use and acceptance of the MDM by the larger marine GIS user community. Testing the MDM was initially done by testing how well the MDM met some of the goals outlined by the MDM Working Group (Wright et al., 2001), including

1. *Provide a common geodatabase structure for assembling, managing, storing and querying marine data in ArcGIS.* Presently, the amount of available GIS data relating to the marine environment is rapidly increasing, and a better way to manage and store marine data is needed. Additionally, the need and desire to share data between and within organizations is also increasing. In response to these needs, the MDM geodatabase provides a structure for assembling, managing, storing and querying marine data in ArcGIS. This means that data can: (a) be stored and managed in one

central database, (b) compiled from various data types/formats into the geodatabase format which is easily shared and exchanged between groups, and (c) easily managed due to the hierarchical organization of the geodatabase.

2. *Provide a geodatabase template for marine users*. The geodatabase template is designed to meet the specific needs of the marine community through the creation of feature classes and tables tailored to specific types of marine data. Thus, marine users can immediately begin to populate the geodatabase rather than designing it from scratch. In addition, the geodatabase provides a mechanism for the implementation of data content standards, in the form of metadata templates. The ArcGIS data model also supports existing data standards, to help simplify the integration of data at various jurisdictional levels (i.e., local, state/provincial, national, global).

3. *Advance data modeling capabilities using object-orientation.* One of the key advantages of the MDM is that it helps users take advantage of the most advanced analysis capabilities of ArcGIS, particularly the ability to capture the behavior of real-world objects in a geodatabase, and the support of more complex rules that can be built into its geodatabases Thus, users can characterize features more naturally by letting them define their own objects, by defining topological, spatial, and general relationships, and by capturing how these objects interact with other objects (Zeiler, 1999). In this regard, the geodatabase model is one step beyond the coverage model (a model without object-orientation capabilities) and one step closer towards a more realistic model of the world. For example, in the coverage model point, line and polygon features have the same behavior; thus, the behavior of a line representing a road is identical to the behavior of a line representing a stream (Zeiler, 1999).

However, through object-orientation, behaviors can be attributed to the stream or road to distinguish between the two features. Additionally, object-orientated data modeling data allows for more accurate data entry and editing due to the ability to establish validation rules for the data.

4. *Improve spatial data representation*. Given the spatial and temporal variability of marine data, the two-dimensional (2D) mapping of the traditional GIS is no longer adequate for marine analyzes (e.g., Wright and Bartlett, 2000; Wright, 2002). Take the example of parameters such as temperature, salinity and sound velocity, which are all measured as a function of depth (the third dimension or 3D) in different locations. The ability to visualize that third dimension of those parameters allows for a more effective perception of spatial relationships. Moreover, 3D visualization systems improve the users' capacity for data mining, as the data are presented in a more intuitive way, often revealing hidden relationships (Kleiner et al., 2000). Thus, the aim of the MDM is to provide a more accurate model of oceanographic phenomena, by improving the representation of marine data.

5. *Increase the understanding of ArcGIS.* Due to the rapid advancement of GIS software, many users are still working in ArcView 3.x. However, the MDM creates a new way to model marine data, and expands marine modeling capabilities and analytical tools, thus providing a reason for people to learn ArcGIS. By working with the MDM, users can become more familiar with the basic ArcGIS layout. In addition, the MDM tutorial (Appendix 1) will aid users in becoming more familiar with GIS in general, and more specifically, with the new geodatabase model.

An essential component of this research, the ArcGIS MDM provides a GIS data structure for assembling, storing, querying and sharing available information on Hawaiian near-shore waters. The information integrated into the MDM during this study will be shared with the Hawaii Natural Heritage Program (HNHP) to assist them in the completion of the Marine Gap Analysis Project (MGAP). The HNHP's MGAP was established for the purpose of integrating available information on Hawaiian near-shore waters in order to assess Hawaii's marine biodiversity, as well as evaluate current MPAs and potentially establish new MPAs (Puniwai et al., 2003). This study is important because comprehensive assessments of the status and trends of biological diversity have historically ignored marine ecosystems (Puniwai et al., 2003), which has led to a poor understanding of current marine biodiversity and a lack of information to serve as the basis for the prioritization for marine conservation areas (Carr et al., 2003). Additionally, many areas were protected based upon the need to manage user conflicts and safeguard protected species, rather than any biological selection criteria (Friedlander and Brown, 2003). The information this study will provide to the HNHP includes the collection, integration and mapping of spatial information of Hawaii's near-shore reef fish around the main Hawaiian Islands in the form of a geodatabase, and specific geographic information system (GIS) layers such as 10-m gridded bathymetry and rugosity data, as well as general habitat information. Not only will this information be of use to the HNHP, but can also serve to increase the efficacy of future monitoring programs, as it synthesizes information from existing monitoring programs in Hawaii and identifies gaps in information.

Materials and Methods

Part 1. Testing the ArcGIS Marine Data Model

Fish transect survey data from Hawaii were obtained from various federal,

local and academic institutions through the HNHP. The data were gathered and

organized in a Microsoft Access relational database which was used primarily for its

querying abilities and large data storage capacity. The data gathered were from:

WHAP, the Coral Reef Assessment and Monitoring Program (CRAMP), the state of

Hawaii's Division of Aquatic Resources (DAR), the Saving Maui's Reefs project and

individual peer-reviewed journal articles (Table 1). As anticipated, the data were in

several different formats, with some data in electronic format and some in reports.

Table 1. The dates, sites and survey methodologies of fish survey data used in each
study. Note: WHAP and Saving Maui's Reef projects also focus on obtaining coral
information; however, this study excluded that information from the database. Also,
information on sessile organisms from the CRAMP data set was included in this
database because this information will be later used by the HNHP in their larger
MGAP project.

Study	Date	No. of Sites	Method Used	Method Details	Frequency	Variables Observed
WHAP (Tissot et al., 2004)	1999-2002	23	Strip Method	Two pairs of divers, each pair surveys two 25-m transects separated by a 10-m buffer.	Six surveys of each site per year	Fish and coral species and abundance
CRAMP (Brown et al., 2004)	1999-Present	30	Strip Method	Digital video of ten 10-m fixed transects to analyze coral cover and growth, and five fixed photoquadrats to examine the	Annually	Coral cover and growth, and the recruitment and mortality of sessile benthic organisms

				recruitment and mortality of sessile benthic organisms		
DAR (DAR, 2004)	1986-1999	28	Strip Method	Varies from four 25-m transects or one 100-m transect.	Varies	Fish species and abundance
Saving Maui's Reefs (Brown, 1998)	1991-1998	13	Strip Method	A pair of Earth watch volunteers survey a 50-m transect, documenting fish and coral species and abundance every 5-m	In 1998, a total of 229 transects were conducted for the thirteen sites (no. of surveys at each site varied)	Fish and coral species and abundance
HNHP*	1971-1983	27	Strip Method	Varies	Varies	Fish species and abundance

*Contact the HNHP for specific metadata information

The data sets obtained through the HNHP contains information on 121 sites located throughout the main Hawaiian Islands (Figure 4). These sites are located in 15 different habitat types and 5 zones, and are located in areas with different levels of protected status (NOAA, 1999a).
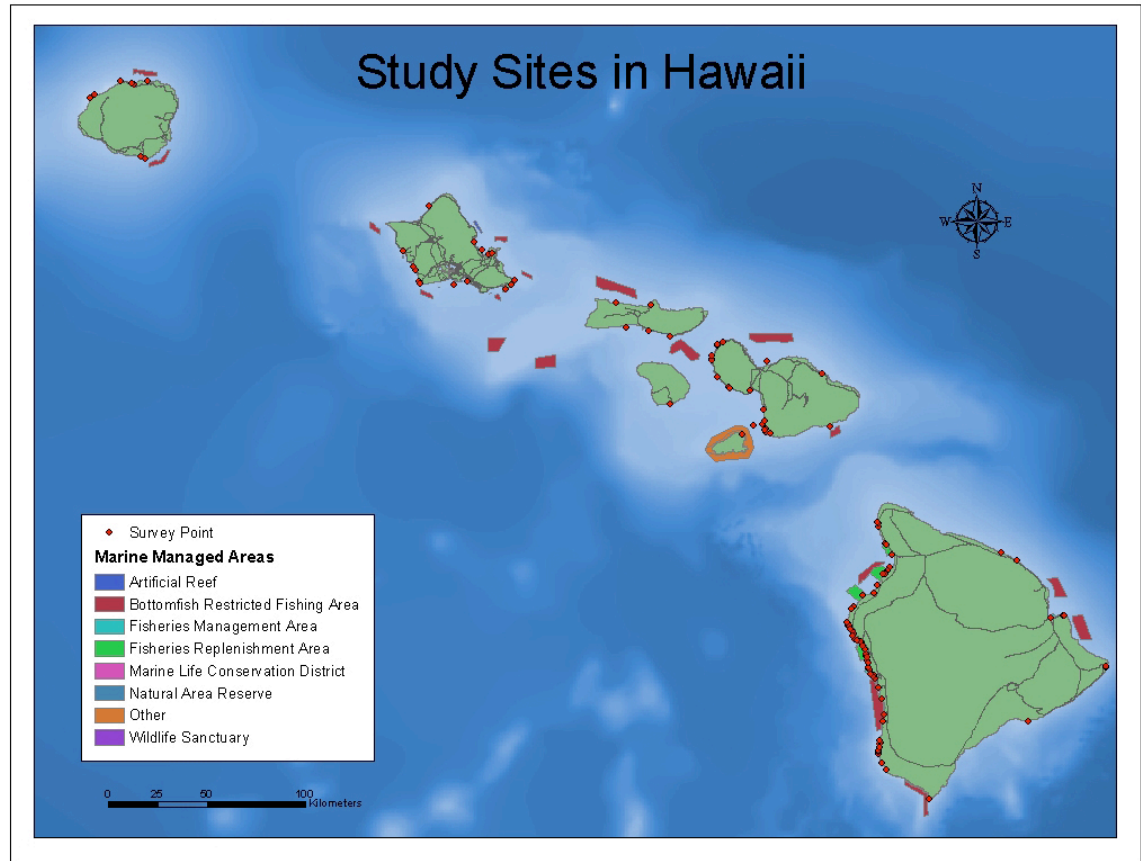
Figure 4. Study sites for data used in the ArcGIS MDM.

The area examined in the habitat utilization analysis is the West Coast of Hawaii Island (see Figure 5). From Lapakahi to Manuka this study represents approximately 212 km^2 of shoreline, fifteen different habitat types and five reef zones. This area was chosen due to the availability of fish transect survey data and habitat information.
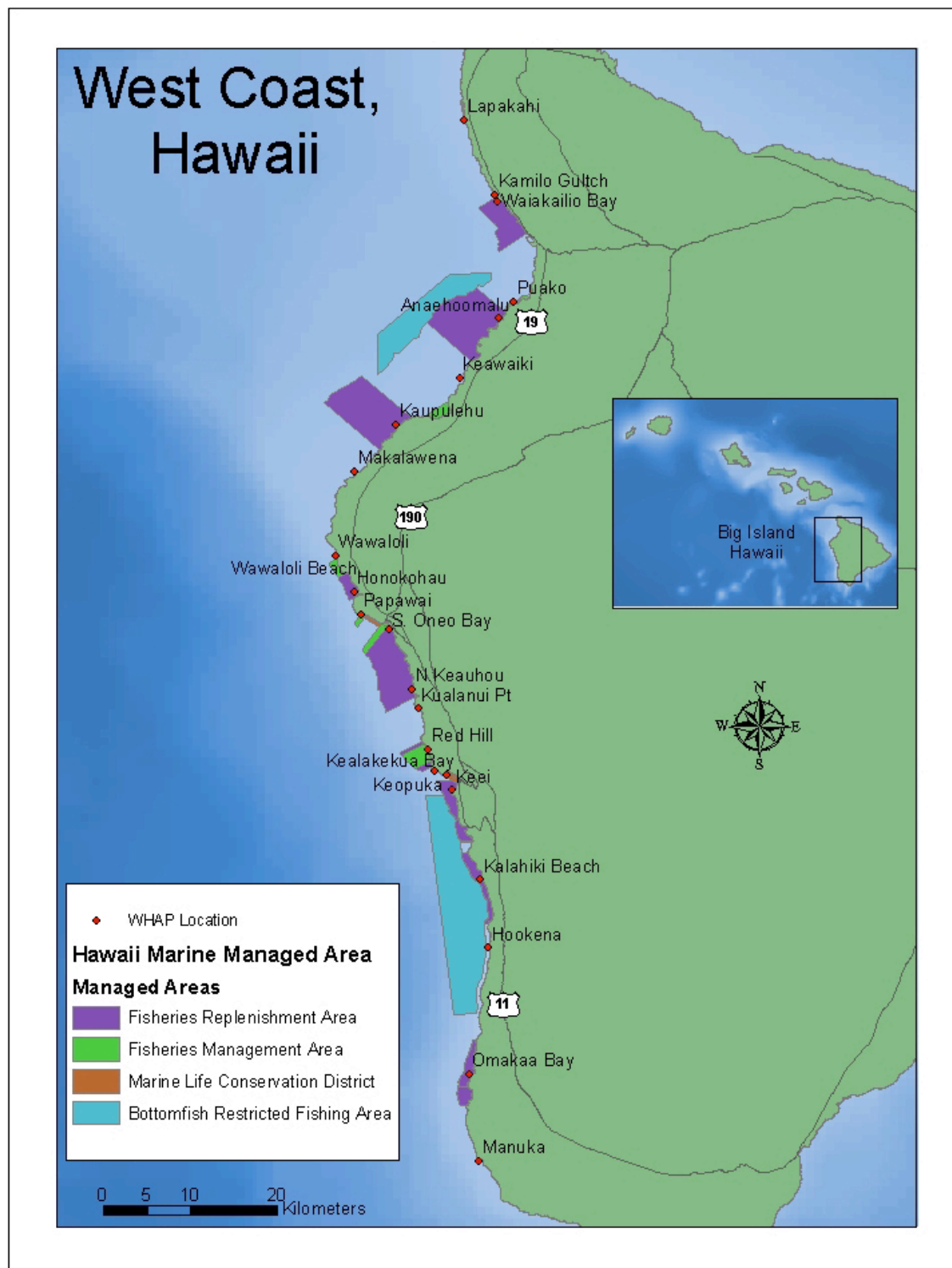
Figure 5. WHAP study sites in West Hawaii.

Data on fishes from these studies were collected using the strip transect method (Hallacher and Tissot, 1999) to record the number of observed fish species, count and estimate their size along 100m^2 transects. The strip method was used to estimate the densities of fish because it provides information on species presence, as well as information on their density (Hallacher and Tissot, 1999). While this method has its biases, some of which include the underestimation of individuals present, missed or misidentified fish species, incorrect estimations of fish abundance, or over-representation of mobile species, it is the most widely used method to estimate fish abundance (Hallacher and Tissot, 1999; Tissot et al., 2004). Therefore, when using the MDM the emphasis will be placed on the observed fish patterns at the selected sites over time.

The WHAP data (Table 2) were gathered from twenty-three permanent locations, collected from 1999-2002. The WHAP surveys were conducted to study the effect of aquarium collecting on natural populations around the Big Island, and to determine the effectiveness of the existing network of MPAs along the West Coast of Hawaii (Tissot et al., 2004). The Saving Maui's Reefs data were gathered from thirteen permanent locations, from 1991-1998. Saving Maui's Reefs surveys were conducted in an effort to detect spatial and temporal changes in the structure of the coral reef community, coral coverage and reef fish density and diversity (Brown, 1998).

Table 2. Data stored in the WHAP and Saving Maui's Reefs database. Data types in bold serve as key fields used to relate one table to another.

Table Names	Divemaster	TransectRun	RunDetail	Species	Location
Information stored in each table	**Survey** Date **Location** Time	**Survey** **TransectRunID** Transect QLocation Observer Run/ Quadrat	ID **TransectRun ID** **Species Code** Number Sum of run P logP	**Species Code** Common Name Species Name Two letter code Trophic Modility Endemic	**Location** Dive site Latitude

The CRAMP data (Table 3) were gathered from thirty locations, collected from 1999-Present. The CRAMP surveys were conducted to quantify changes that have occurred on coral reefs subjected to varying degrees of fishing pressure, sedimentation, eutrophication, and algal growth, and to test hypotheses concerning the role of these environmental factors in the ecology of coral reefs (Brown et al., 2004).

Table 3. Data stored in the CRAMP database.

Table Names	CRAMP_Sites	CrampFishTotal
Information stored in each table	**SiteID** Site Name Island Latitude Longitude Full Latitude Full Longitude Management Status Area Description	ID **SiteID** Depth01 Depth02 Species

The DAR data set (Table 4) represents twenty-eight locations around the main Hawaiian Islands, from 1986-1998. The methodology behind the collection of these data are uncertain. This data was collected by DAR to provide baseline information on fish populations around the state of Hawaii in order to better manage the fisheries.

Table 4. Data stored in the DAR database.

Table Names	Sites	Surveys_Revised	SurveyData
Information stored in each table	ID **Site Code** Area Island DAR variants Comments	ID **Survey Code** **Site Code** Month Day Year Date Survey Area Original Dar Files Notes Source	ID **Survey Code** Species Code Count Size (in)

The HNHP database is a compilation of data sets gathered from many peer-reviewed journal articles (Table 5). This database covers twenty-seven sites, collected from 1971-1983. All data were collected visually; along transect survey lines ranging from 25-m to 100-m. For detailed information on survey methodology and the purpose of each study contact the HNHP.

Table 5. Data stored in the HNHP database.

Table Names	Divemaster	TransectRuns	RunDetail	Observers
Information stored in each table	**SurveyID** Date Time Location	**SurveyID** **TransectRunID** Transect **ObserverID** Run/Quadrat	**TransectRunID** TaxonID Count Size Comments	**ObserverID** Observer name

All datasets were combined into a single MS Access database created to store these data. This database consisted of four primary tables: (1) survey location information, (2) Divemaster (3) RunDetail and (4) species information (Table 6 for specific information). All tables had a key field that was used to relate one table to another, allowing the user to query related information throughout the tables. This database design allows for queries involving variables of space, time, and fish species.

Table 6. The database structure of the combined Hawaii reef fish database.

Table Names	Survey Location (Survey Point)	Divemaster	RunDetail	Species Information (Integrated Species)
Information stored in each table	SurveyID **Reference location** Latitude Longitude Island Site Area_Desc Depth Comments	**Reference location** **Transect_Co** TransectRunID Transect Date Time Observer Source	**Transect_Co** **TaxonID** Count Size Comments	**TaxonID** EBrownID TaxonName Alpha Type Family FamCode Quest DAR Synomyn HawaiiName CommonName Trophic Status

Testing goals of the MDM:

Goals 1 and 2*: Provide a Common Structure for Assembling, Managing, and Querying Marine Data in ArcGIS and Provide a Geodatabase Template for Marine Users*

In the MDM, the feature data sets and classes cover a very broad range of marine applications (Wright and Blongewicz, 2003) (Figures 6 and 7). For example, the marine feature data set contains seven feature classes used to store information representing physical maritime features such as natural or manmade objects in the water.

Common Marine Data Types

Draft 6/28/03 by Pat/Dawn

Feature Points
Previously "Fixed Points"

Measurement Points

Feature Lines

Survey Point

ID
X,Y
Z

Examples:
marker buoy,
transponder,
other fixed
geography

Instantaneous Point

ID
X,Y,
Z or ΔZ
t,
$m_1...m_2$

Examples:
CTD, XBT, SVP
casts at ΔZ, fish
density, tide
gauge, etc. at
surface or a single
Z

Location Series Point

ID
$\Delta X, Y$,
ΔZ
$m_1...m_2$
$t_1...t_2$

Examples:
telemetry,
bird/mammal
sighting, ship
mounted
ADCP

Time Series Point

ID
fixed X,Y,
Z or ΔZ
$m_1...m_2$
$t_1...t_{infinity}$

Examples:
current meter,
moored ADCP at
ΔZ, obs. buoy,
hydrophone,
OBS at single Z

Data Line

ID
X, Y
M_1, M_2
$Z_1, Z_2...$
Examples:
for abstracting data from,
building profile from,
measuring change along...
**cross-section,
bathy profile,
sed transport line**
where "M" is GIS geometry
measure, rather than a
"measurement" from an
instrument in the field ("m")

Time Duration Line

ID
X1,Y1,X2,Y2...
$M1, M_2...$
Z1,Z2...
m1,m2...
t1,t2...

Examples:
algal bloom trawl,
ADCP tracks,
ARGO drifter

Run **Track**

Feature Line

ID
$X_1, Y_1, X_2, Y_2...$
$M_1, M_2...$
Examples:
sea wall, ocean front,
EEZ or
legal boundaries NOT
enclosing an area

Shoreline

shoreline type,
VDatum

Marine Areas

Derived or Placeholders

Feature Area

ID
$X_1, Y_1, X_2, Y_2...X_1, Y_1$
Z
m

Examples:
Marine boundaries
(e.g.,sanctuary, MPA),
habitats,
patches, lava
flows, clipping,
masking

Time Duration Area

ID
$X_1, Y_1, X_2, Y_2...X_1, Y_1$
Z
m
$t_1...t_n$

Examples:
No-take
zones,
oil spills,
harmful algal
bloom

Regurlarly Interpolated Surface

$row_1, col_1...row_n, col_n$
$Z_{r,c1}..Z_{r,cn}$

Examples:
bathymetry,
sidescan,
SST, climatology,
"re-analyzed"
products
(*placeholder for*
IMAGES such as
GeoTIFF, BIL, etc.)

Irregularly Interpolated Surface

$row_1, col_1...row_n, col_n$
$Z_{r,c1}..Z_{r,cn}$

Examples:
TINs
bathymetry

Volume

ID
$X_1, Y_1, Z_1, X_2, Y_2, Z_2$
$...X_1, Y_1, Z_1$
m or $m_1...m_n$
t or $t_1...t_n$
ncols,nrows,nlayers

Examples:
plume, front,
warm core, trawl
abundance

Animations, Movies, Video

$X_1, Y_1, Z_1, X_2, Y_2, Z_2$
$...X_1, Y_1, Z_1$
$t_1...t_n$

Examples:
U/W video footage,
outputs from atm. or
circulation models
that are animated &
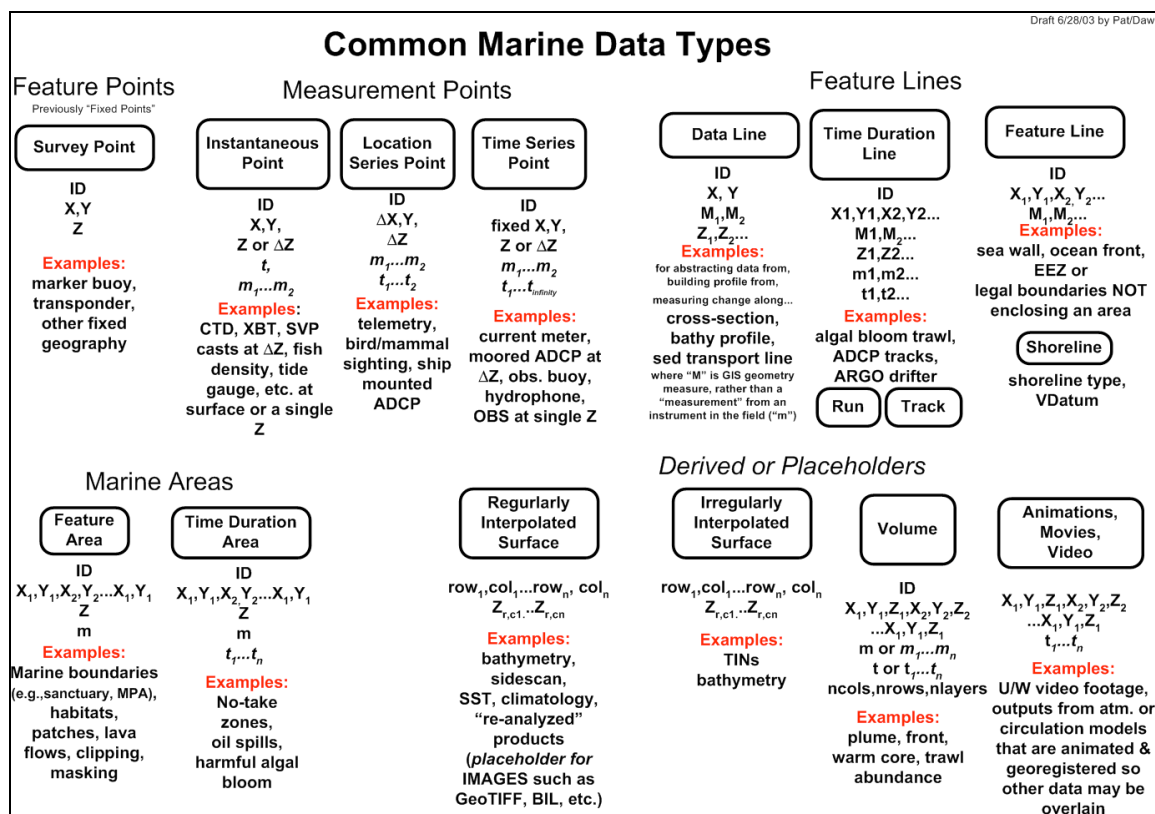georegistered so
other data may be
overlain

Figure 6. The different marine feature classes and associated data types that can be input into the MDM (Wright and Blongewicz, 2003).
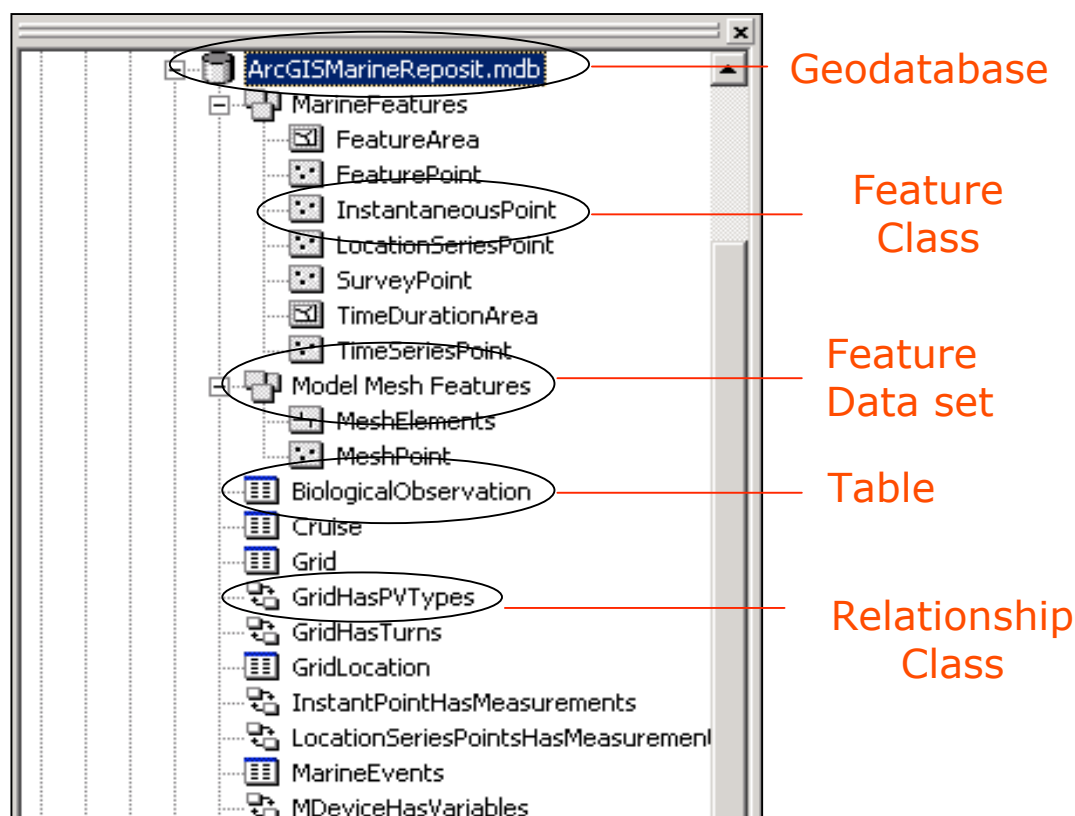
Figure 7. A screen snapshot from a directory listing in ArcCatalog showing a portion of the MDM repository.

One feature class within the marine feature dataset is SurveyPoint. The

SurveyPoint feature class of the MDM was created to store information on features

that have fixed x, y and z coordinates, and was used in this study to store the spatial

data integrated into the model. Accordingly, each point in the SurveyPoint feature

class represents a 3D location where a transect survey was conducted. It should also

be noted that many point locations represent only an approximation of the location

where the fish transect survey was conducted, as the exact latitude and longitude

coordinates are unknown. However, the information recorded in the geodatabase

reflects the best possible estimate.

Before any data were added to the geodatabase, the spatial extent and the

projection of the feature classes were established. This was done by importing the

NAD83 UTM Zone 4 projection information and spatial extent from a polygon shapefile of the main Hawaii islands (DBEDT, 2004) into the MDM. Accordingly, all subsequent shapefiles imported into the geodatabase would need to be in the same projection (ie, NAD83 UTM Zone 4), and fall within maximum x, y spatial extent.

The Access table containing the transect survey latitude and longitude coordinates was then imported into ArcGIS 8.3. Using the create x, y function, a point shapefile was generated. This 2D shapefile was converted into a 3D shapefile using the z values (depth) in the attribute table. This was necessary because all shapefiles imported into the SurveyPoint feature class must be 3D in accordance with the established model parameters, even if the 3D values do not exist.

It was initially thought that the SurveyPoint feature class could be further defined with information stored in the SurveyInfo object class. However, due to the large amount of data, this did not prove to be possible.   With over 250,000 entries, if all of the data was combined into one table representing the main Hawaiian islands, it would take the MDM over 2 hours to perform one query. Thus, by separating data by island, queries could be conducted in under one minute, and multiple queries could be conducted if information on more than one island was needed. As a result, the MDM was personalized to fit the data in this analysis. To personalize the MDM numerous field names (Figure 8) and tables were added to the geodatabase. The added tables were: Divemaster, RunDetail and Integrated Species, which were specific to each island (i.e., Oahu Divemaster, Hawaii Rundetail, ect). These tables further describe the SurveyPoint feature class by providing information like survey dates, who conducted the survey, and which species were observed along transects.
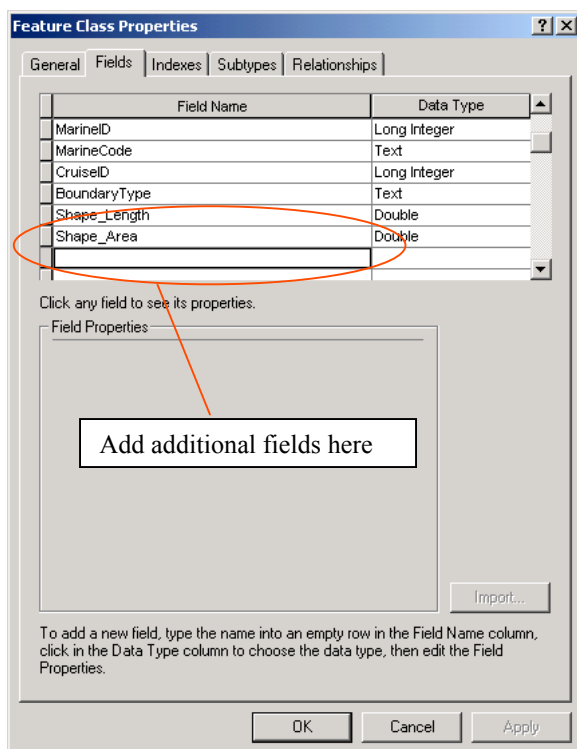
Figure 8. The MDM feature classes were personalized by adding additional field names.

When all the data were entered into the geodatabase, relationships were established between the feature classes, or spatial information, and the tables, or non-spatial information (Figure 9). These relationships consisted of one-to-one and one-to-many relationships. A one-to-one relationship matches one entry to an identical entry in a separate column/table. A one-to-many relationship matches one entry to multiple identical entries in a separate column/table. Additionally, due to the nature of the geodatabase, these relationships are permanent and unlike joins, will not have to be reestablished in each new project.
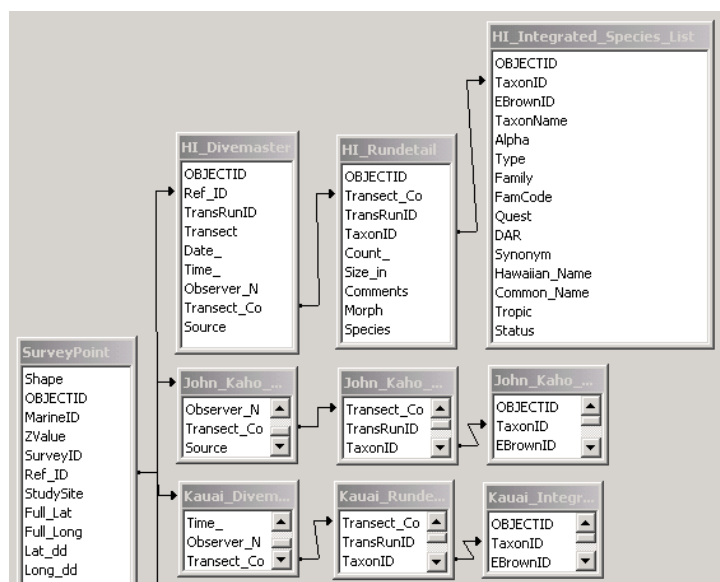
Figure 9. The relationships established between the tables, as seen in MS Access.

One of the main reasons the MDM was used in the habitat utilization analysis

was due to the ability of the model to easily and efficiently query for spatial

information. Accordingly, the established database design allows for queries

involving variables of space, time, and fish species. Thus, by querying the MDM, the

site locations where a specific fish species has been observed (Figure 10) can be

determined. These sites are then overlaid on large-scale habitat types to determine

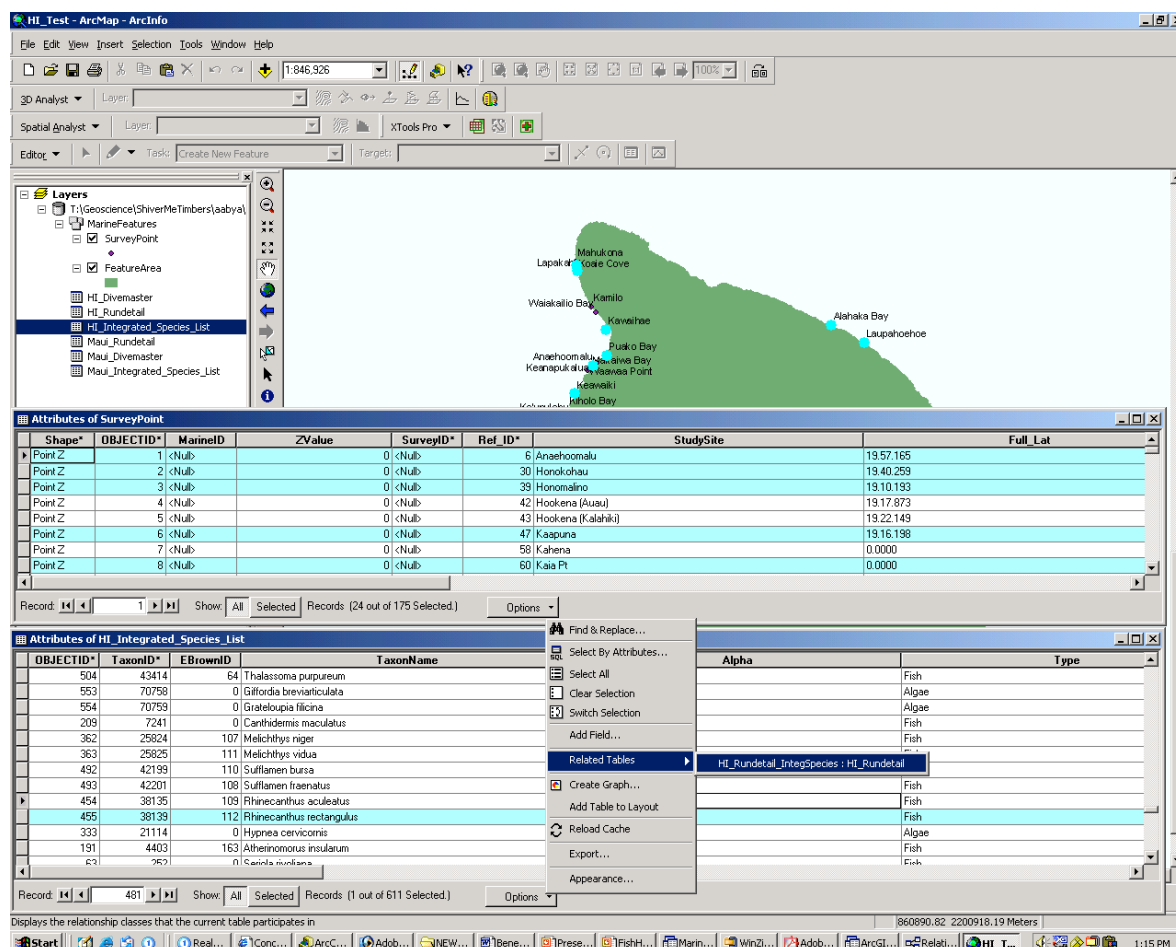what species were found in what habitat types.

Figure 10. Illustration of a query in ArcGIS using the MDM to show the sites where the *Rhinecanthus aculeatus* (Blackbar triggerfish) has been observed. This information was used in the regional habitat analysis.

Goal 3. *Advance the data modeling capabilities using object-orientation*

Object orientation, or the ability to assign validation rules and behaviors to

data, is one of the main reasons why people are using geodatabase models:

> The defining purpose of [the geodatabase] is to let you make the features
> in your GIS data sets smarter by endowing them with natural behaviors,
> and to allow any sort of relationship to be defined among features
> (Zeiler, 1999).

However, I did not attribute any behavior to my data in this case study. This was

because object orientation was not needed to answer the question of habitat utilization

by reef fish species. It would have been possible, if time had allowed, to assign

validation rules to the fish sizes in the MDM data set. This would have been done by

assigning a size range for each different fish species at which it is known to exist.

This would have highlighted possible errors in the data sets if fish sizes fell outside of

the known possible size range.

Goal 4. *Improve Spatial Data Representation*

To test whether or not the MDM improved spatial data representation, different

types of data were represented in ArcScene, which is a 2.5D viewing program. The

2.5D is different than 2D in that it includes a third-dimension, but is not quite 3D in

that the depth or elevation is an attribute of a point, line, or area, rather than being a

part of the location in the data structure. In doing this, the ease with which data could

be viewed in ArcScene was tested, along with how well the data could be spatially

visualized in 2.5D was observed. First, the SurveyPoint feature class was mapped in

ArcScene using the depth attribute values. And second, the bathymetry and rugosity

grid layers were visually represented in ArcScene using the depth values to create a

2.5D representation.

Goal 5. *Increase the Understanding of ArcGIS*

Due to the rapid development of GIS software, many users are still working in

ArcView. To help users become more familiar with GIS in general, and more

specifically with the geodatabase model, a tutorial describing how to use the MDM

was created as a part of this research. My colleagues who have used the MDM tutorial

describe it as easy to use and understand. In addition, they state that they like the

diagrams and the formatting.

Part 2: Habitat Analysis Methodology

Once the near-shore reef fish data sets were imported into the MDM, the spatial habitat utilization patterns of reef-fish along West Hawaii Coast were examined. To determine these patterns, three spatial data sets were used (1) the West Hawaii Aquarium Project (WHAP) fish survey data set stored in the MDM, (2) NOAA's large-scale habitat delineations stored in the MDM and (3) the WHAP substrate data set. To answer the question of habitat utilization, four sub-questions were investigated: (A) How do specific fish species utilize the RA and RCB habitats? (B) Does depth play a significant role in the classification of RA and RCB habitat types? And (C) Is small-scale substrate information correlated with large-scale NOAA habitat information? (Figure 11).
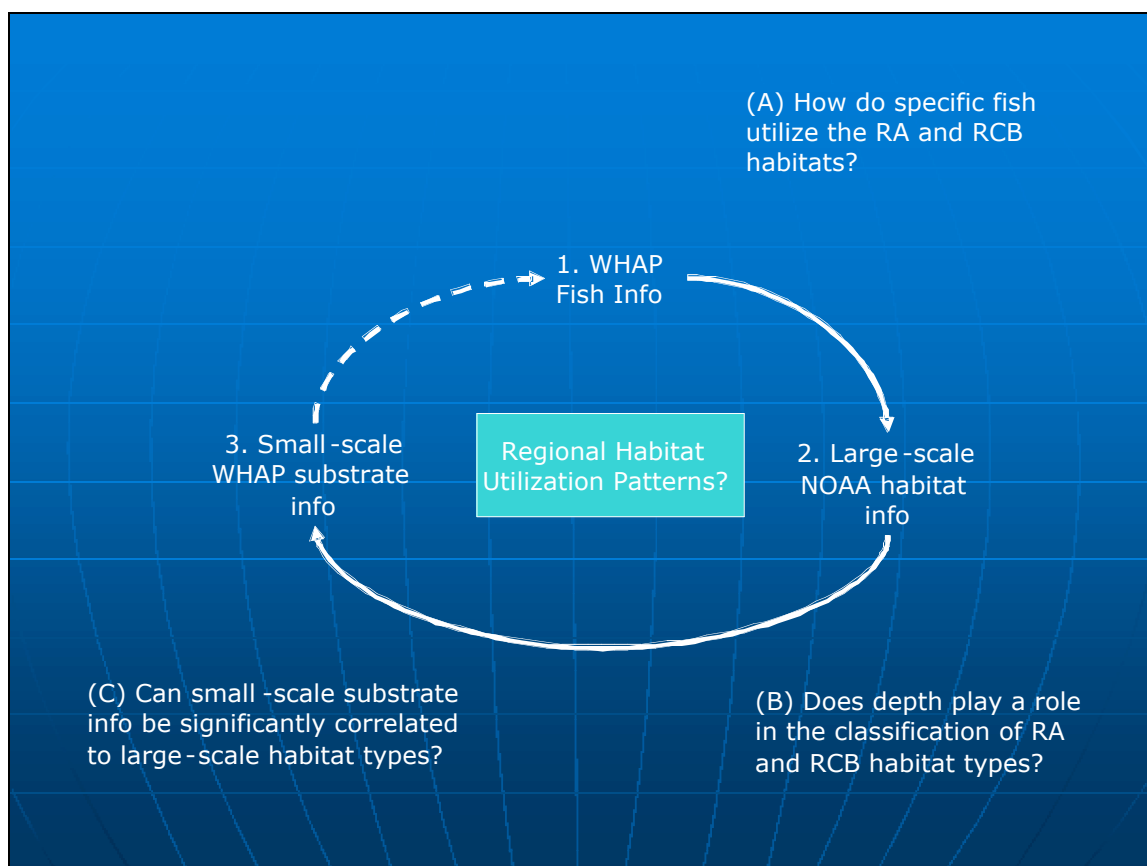
Figure 11. Steps taken to determine regional habitat utilization patterns for selected fish species.

The WHAP data set was used because it is one of the most comprehensive and statistically robust Hawaiian reef fish data sets. Additionally, it covers twenty-three sites along the West Coast and it spans a multi-year time period. The detailed steps taken to answer each of the four sub-questions are outlined below (Figure 12).
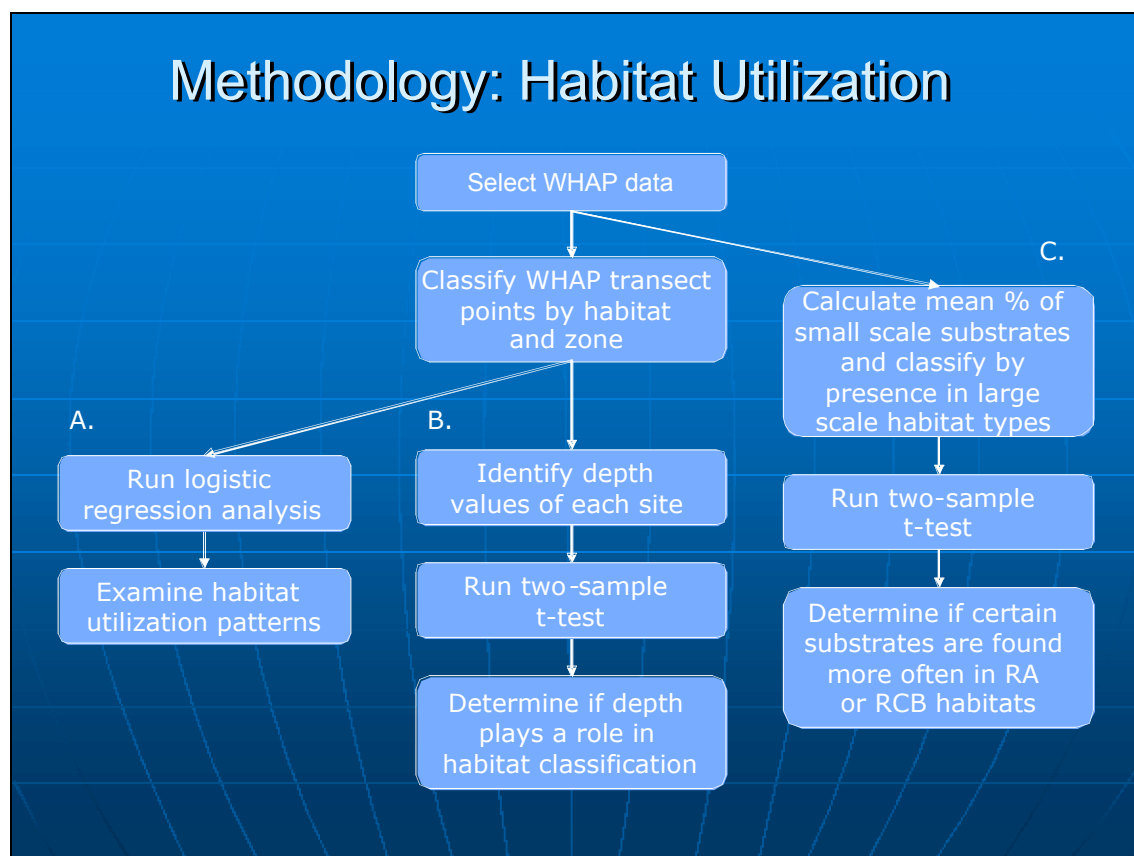
Figure 12. The specific steps taken in the fish analysis portion of this study.

In order to analyze the WHAP data set in this analysis, the MDM was used to

query and separate the WHAP data from the other data in the geodatabase. It was

necessary to isolate data specific to WHAP, so as not to confuse a fish species

observed on a DAR survey at the same site as fish species observed on a WHAP

survey. Moreover, it was important to separate the WHAP data because survey

methodologies differed between data sets, because an analysis of a compilation of data

sets would not be statistically valid. When the data were separated, a new feature

class called Whap Location was created, along with 3 tables: Whap Divemaster, Whap

Rundetail and Whap Species. All of these data remained in the MDM, and new

relationships between the tables and feature class were formed following the initial
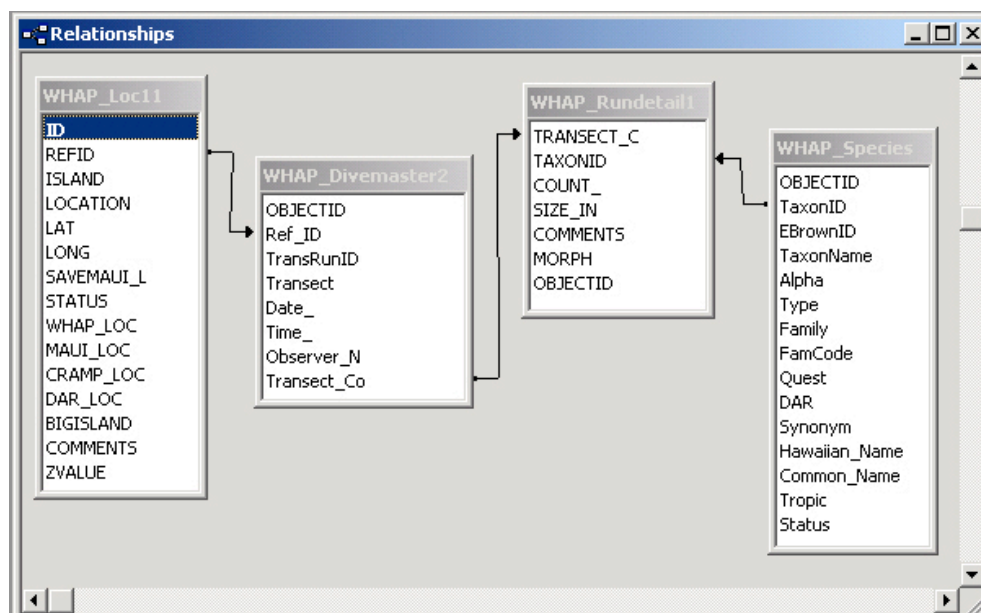
MDM database design (Figure 13).

Figure 13. The WHAP relationships as viewed in MS Access.

The locations of the WHAP transect survey sites were double checked by comparing latitude and longitude coordinates with the current coordinates stored in the WHAP Location feature class. After making minor revisions, these locations were mapped in ArcGIS. To determine the habitat type and zone for each of the WHAP sites, the WHAP Location feature class was overlain on a NOAA's large-scale benthic habitat layer (Figure 14). Using the select-by-location function in ArcGIS, the habitat types and benthic habitat zones for each of the WHAP sites were determined.
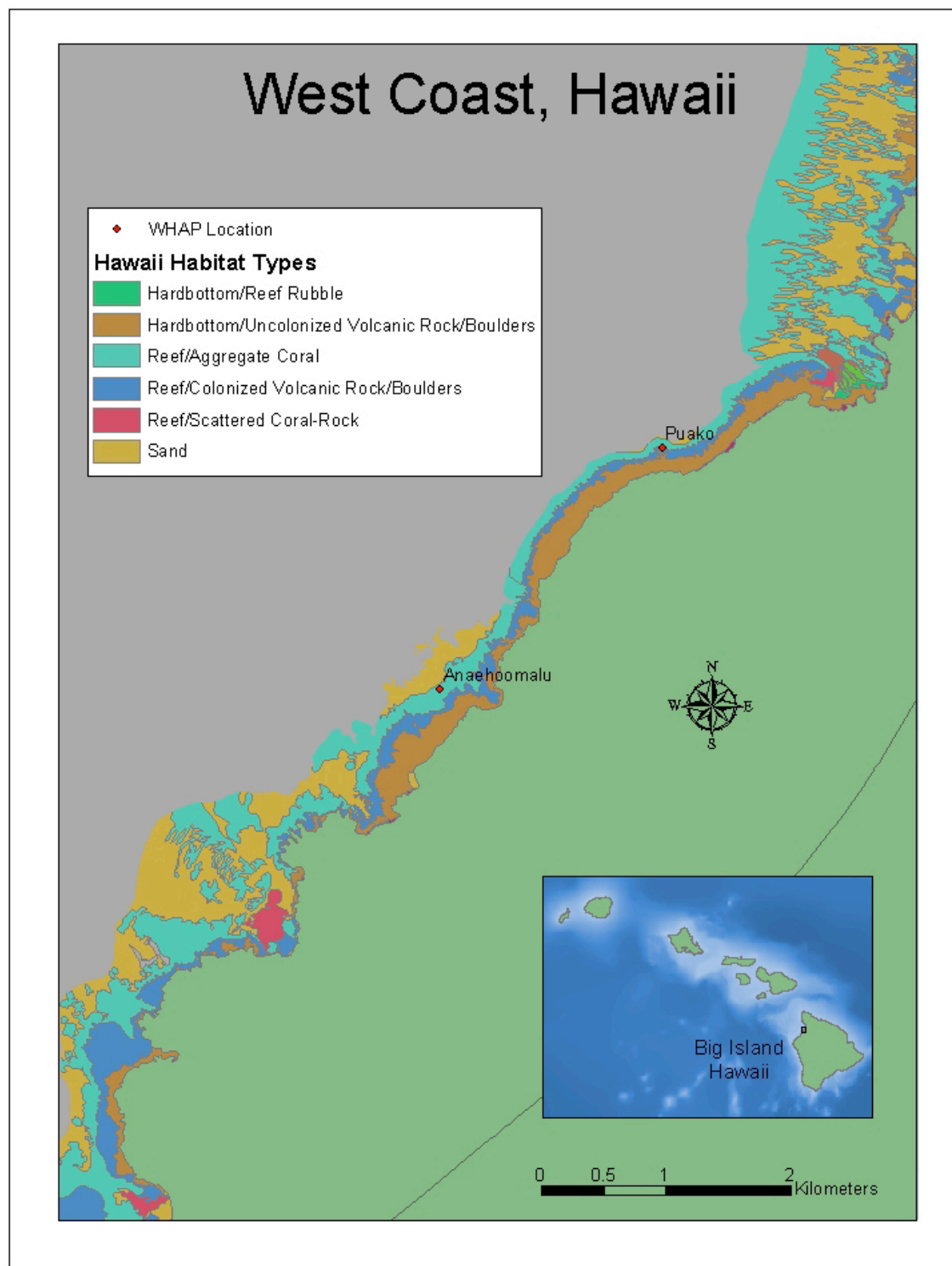
Figure 14. The benthic habitat shapefile overlain with WHAP location points to determine the habitat type in which the sites were located.

Once each WHAP site was classified according to habitat type and zone, a logistic regression analysis was used to determine the pattern of habitat utilization of six aquarium and seven non-aquarium fish species (Table 7) between the different habitat types. These fish species were selected because they represent commonly collected aquarium and reef fish observed during WHAP surveys (Tissot, pers. comm.).

Table 7. Aquarium and non-aquarium fishes examined.

Aquarium Fish	Non-aquarium fish
Acanthurus achilles	*Acanthurus nigrofuscus*
Chaetodon ornatissimus	*Acanthurus triostegus*
Chaetodon quadrimaculatus	*Ctenochaetus strigosus*
Forcipiger spp	*Paracirrhites arcatus*
Naso lituratus	*Paracirrhites foresteri*
Zebrasoma flavescens	*Scarus dubius*
	Scarus psittacus

Following the methodology outlined by Christensen (2003), a logistic regression was used to fit a model to a binary response (Y=1 if present or 0 if not present) to the independent variable (X=habitat type), such that for each column, there was a probably p of being present or p-1 if not being present. The MDM model was used to query the location of each specific fish to determine its presence or absence at each WHAP site.

Further analysis into the habitat utilization of specific fish species was conducted to determine if there were correlations between the small-scale substrate information gathered in 1999 by the WHAP researchers (Tissot et al., 2004) and the large-scale habitat information delineated by NOAA in 2000 (NOAA, 1999a). To do this, the mean percentages of each small-scale substrate type found at each WHAP site

were calculated. The WHAP sites were then categorized according to habitat type,

and the information was compared using a two-sample t-test.

To determine if depth values played a role in the classification of NOAA's

large scale habitat types, the depth values of each of the WHAP sites were determined.

The WHAP sites were then categorized by habitat type and the information was

compared using a two-sample t-test.

Additional Data Sets

Benthic Habitat Layer. Specific habitat types were determined from the NOAA

benthic data shapefile. The line work in this shapefile was generated from remotely

sensed data (hyperspectral and IKONOS satellite imagery) for the Island of Hawaii

during the 2000 acquisition mission (NOAA, 1999a). Based upon this imagery, a

hierarchical classification scheme was created to define and delineate habitats zones

(NOAA, 1999b). This classification shows only continuous habitats greater than 1

acre in size (Battista, 2003).

Rugosity Layer. The rugosity layer was mapped using the Scanning Hydrographic

Operational Airborne Lidar Survey (SHOALS) data from the US Army Engineer

District (SHOALS, 2000).   The SHOALS data were collected in 2000.   The LIDAR

data were used to derive both the bathymetry, as well as the rugosity grid. Both of

these grids were clipped to exclude areas that overlaid land or were located in the

intertidal zone. Rugosity was determined by using the ESRI script, "Surface Areas

and Rations from Elevation Grids." This extension calculates the rugosity by dividing

the seafloor area by the surface area (ie, a value of 1 equals a completely smooth sea

floor). This calculation is much like the chain link method, where the transect chain length at the bottom is divided by the chain length from the surface; however, this method calculates the value for an area as opposed to a single transect line (Jenness and Engelman, 2004).

Results and Discussion

Part 1. Testing the ArcGIS Marine Data Model

During this study, data from WHAP, CRAMP, DAR, Saving Maui's Reefs and individual peer-reviewed journal articles were successfully imported and managed in the MDM. These data were then queried to determine reef fish habitat utilization patterns in West Hawaii. As the MDM provided a common structure for successfully managing, storing and querying data in the MDM, Goal 1 of this study was met.   In addition, this study showed that the MDM provides a geodatabase template that is tailored for marine data, thereby meeting Goal 2. The MDM template was modified to meet the specific needs of this study through the addition of tables, and field names to the feature class. The ease with which the model can be modified will save the user time in the long run, and is likely comparable to the time it would take to set up a new geodatabase from scratch.

In terms of meeting Goal 3, by advancing the data modeling capabilities using object-orientation, I would argue that this goal was not met for the average user, but was met for GIS developers. I question how easily it is going to be for the average user to create behavior rules in a system that is as dynamic as the ocean. Theoretically, it would be nice to have feature classes such as a dynamic shoreline, but how easy will it be for the average user, especially someone new to ArcGIS, to establish the behavioral rules?   Knowing this, should some common marine behaviors like the dynamic shoreline or seasonality already be programmed in the MDM?

To test whether or not the MDM improved spatial data representation, two different views were compared in 2D in ArcMap and 2.5D in ArcScene. The first comparison was of survey points at different depths overlain upon the bathymetry grid data (Figures 15 and 16), and the second comparison was of survey points at different depths overlain upon the rugosity grid data (Figures 17 and 18).
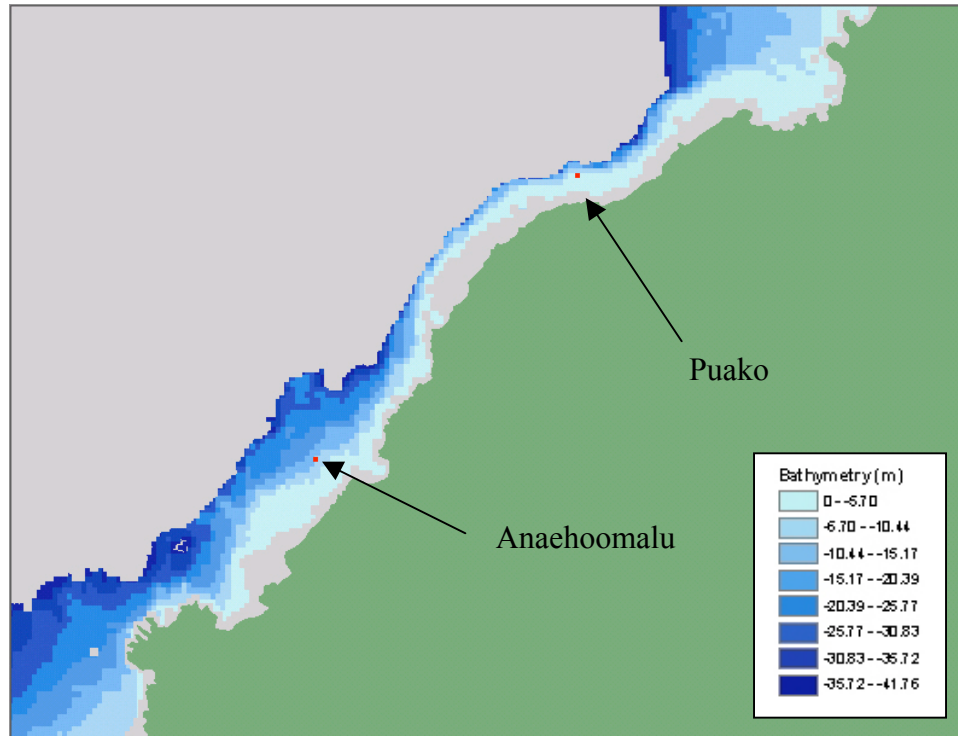

Figure 15. A 2D view of the WHAP Location feature class points overlaid on the bathymetry grid.
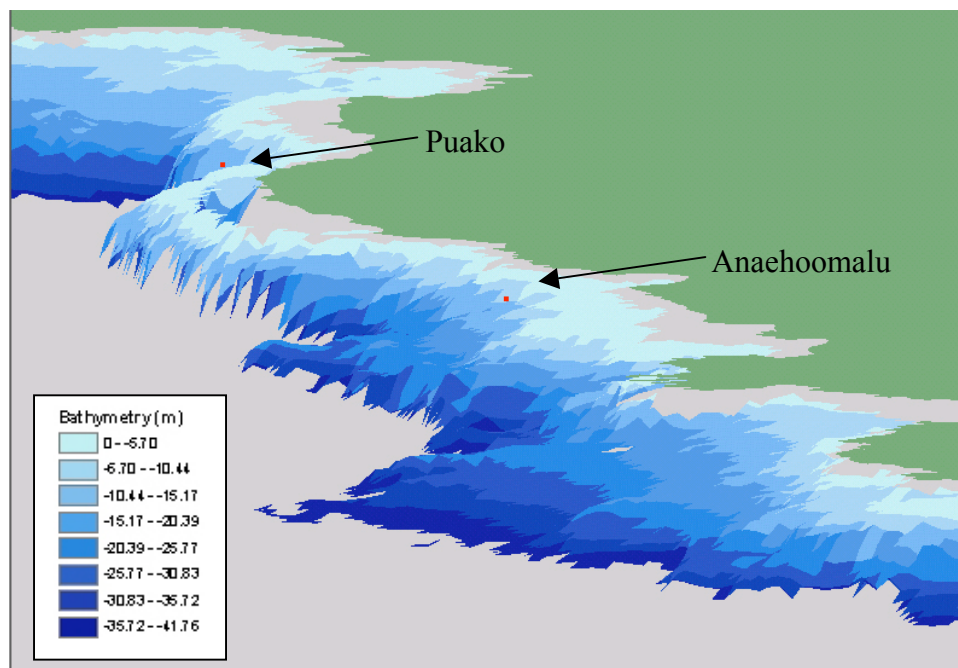
Figure 16. A 2.5D view of the WHAP Location feature class points overlaid on the bathymetry grid.
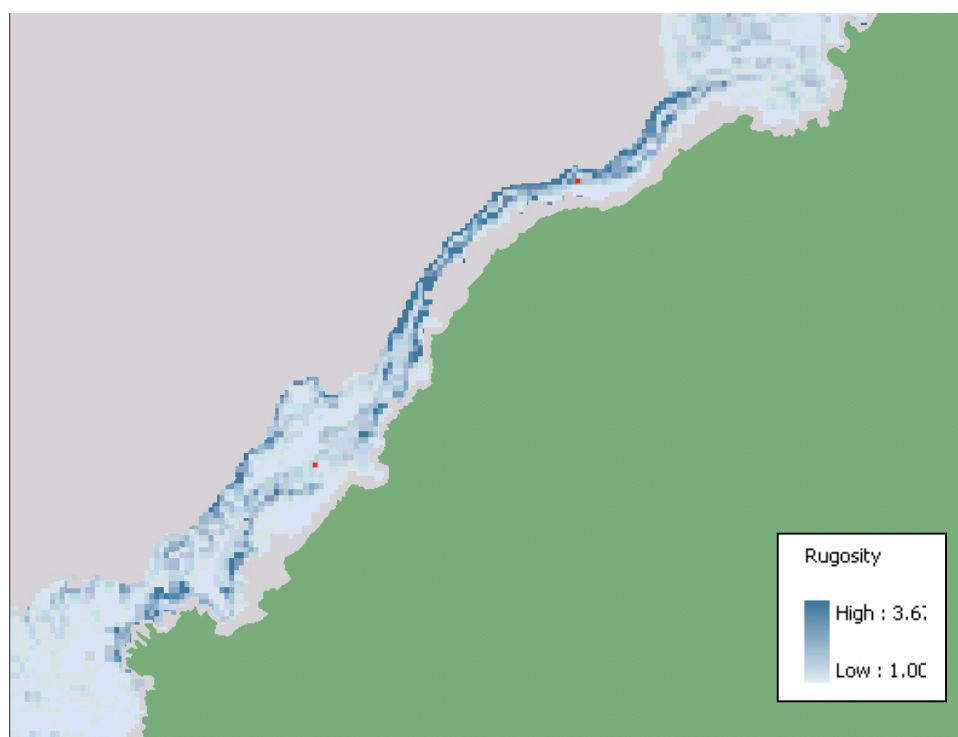

Figure 17. A 2D view of the WHAP Location feature class points overlaid on the rugosity grid.
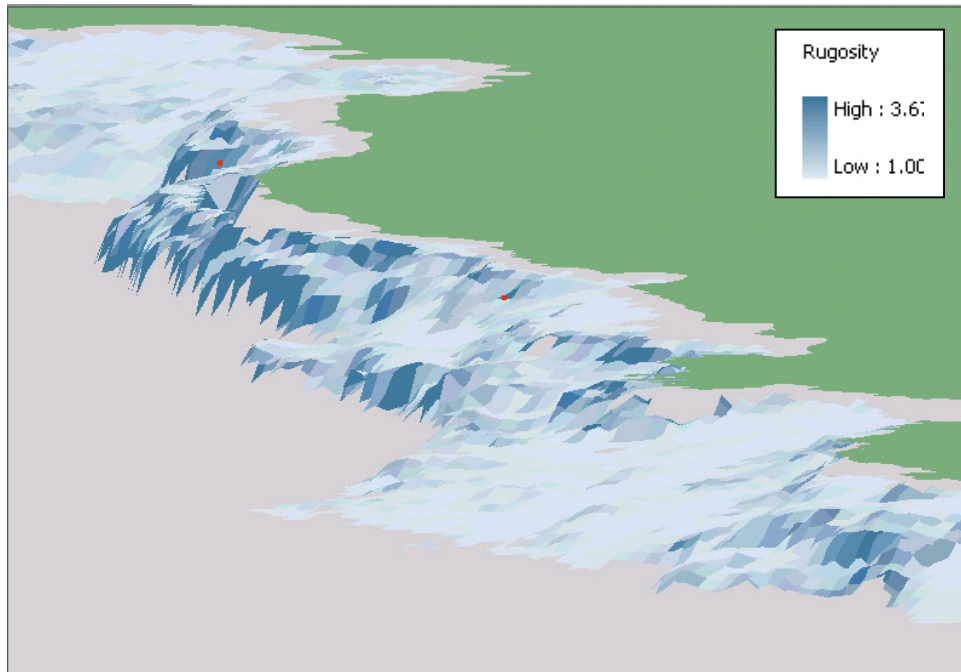
Figure 18. A 2.5D view of the WHAP Location feature class points overlaid on the rugosity grid.

The comparison of the 2D and 2.5D views show that the 2.5D views greatly improved the spatial representation of the data as demonstrated by the better visualization of both the bathymetry and rugosity depth properties in the 2.5D views. Moreover, viewing the data in 2.5D is particularly helpful when examining the location of the survey points in relation to the bathymetry and rugosity layers. In the 2D view, the location of the Puako survey site on the edge of a drop-off is not as apparent as it is in the 2.5D view. This visualization can aid greatly in understanding the ecological differences in flora and fauna found at Puako vs. a site like Anaehoomalu which is located on a gentler slope. Yet, it should be noted that the processing of grids and their spatial representation is the same whether or not one is using the MDM.   Additionally, the author found that some of the useful functions, like labeling points, were not available in ArcScene.

The last goal in the design of the MDM is to increase the understanding of ArcGIS and geodatabases in general. The results of this analysis have shown that in conjunction with the tutorial, this goal has been met. My colleagues at Oregon State University who have used the tutorial to become more familiar with the MDM state that it is easy to use and understand. Also, they state that they have become more familiar with tool placement and analytical capabilities in ArcGIS. In addition, for those users with very little geodatabase experience, working through the tutorial will familiarize them with terms and concepts associated with geodatabases, as well as show them how to work with geodatabases.

Part 2: Habitat Analysis

For the twenty-three WHAP fish sites, the benthic habitat type and zone of each WHAP transect survey site was determined (Tables 8 and 9) through the use of a GIS, by overlaying the WHAP location feature class on NOAA's benthic habitat shapefile.

Table 8. WHAP study sites categorized by habitat type.

Habitat Type	Reef/Aggregate Coral	Reef/Colonized Volcanic Rock/Boulder	No data
Location	Anaehoomalu	Kalahiki Beach	Honokohau
	Keawaiki	Kamilo Gultch	Hookena (Auau)
	Keei	Kaupulehu	Makalawena
	N. Keauhou	Kealakekua Bay	Manuka
	S. Oneo Bay	Keopuka	Omakaa
	Red Hill	Kualanui Pt	Wawaloli
	Waiakailio Bay	Lapakahi	Wawaloli Beach
		Papawai	
		Puako	

Table 9. WHAP study sites categorized by reef zone.

Zone	*Fore-Reef*	*Shelf*	*No data*
Location	Puako	Anaehoomalu	Honokohau
		Kalahiki Beach	Hookena (Auau)
		Kamilo Gultch	Makalawena
		Kaupulehu	Manuka
		Kealakekua Bay	Omakaa
		Keawaiki	Wawaloli
		Keei	Wawaloli Beach
		Keopuka	
		Kualanui Pt	
		Lapakahi	
		N. Keauhou	
		Papawai	
		Red Hill	
		S. Oneo Bay	
		Waiakailio Bay	

From these tables, it can be seen that seven sites are located in the RA habitat type, nine sites are located in the RCB habitat type, and seven are unknown. The RA habitat type is defined as, "coral dominated formations with high relief and structural complexity. [They] often serve the same role as linear reef in fringing reef systems where the reef crest is relatively unorganized" (Battista, 2003), (see Figure 19).

Figure 19. Photograph (Battista, 2003) of a typical Reef/Aggregate Coral (RA) habitat type.

The RCB habitat type is defined by NOAA as "solid volcanic rock that has coverage of macroalgae, hard coral, zoanthids, and other sessile invertebrates that begins to obscure the underlying surface" (Battista, 2003) (see Figure 20).
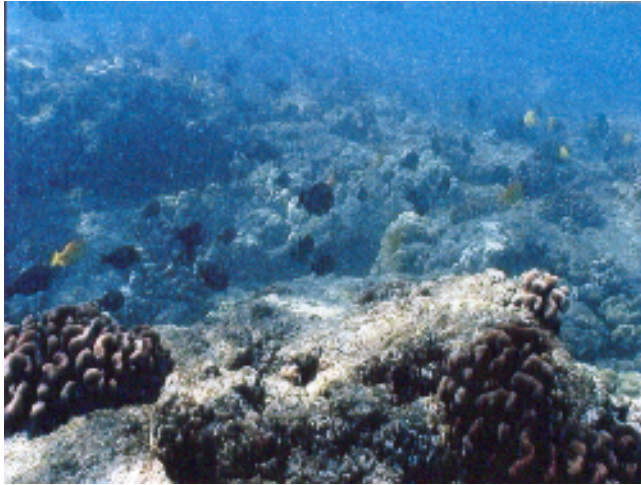

Figure 20. Photograph (Battista, 2003) of a typical Reef/Colonized volcanic rock/Boulder (RCB) habitat type.

The seven 'no data' sites do not have a habitat or zone classification because NOAA's benthic habitat layer does not cover all areas along the West Hawaii Coast. Additionally, it can be seen that all classified site locations, except for Puako, are located in the shelf zone or are unknown. As a result, this habitat characteristic is not used any further in the analysis.

Once the WHAP sites were categorized by habitat type, a logistic regression was used to determine the habitat utilization patterns of the selected reef fishes. It was found that eight fish species: *Zebrasoma flavescens, Chaetodon ornatissimus, Forcipiger spp., Acanthurus achilles, Naso lituratus, Paracirrhites arcatus, Paracirrhites foresteri, and Acanthurus nigrofuscus,* were found 100% of the time in both habitats, *Chaetodon quadrimaculatus* was absent in four out the seven RA sites, but found 100% in the RCB sites, *Scarus dubius* and *Ctenochaetous strigosus* were

absent in one out of the seven RA sites, but found 100% in the RCB sites, and

Acanthurus triostegus was absent four out of the seven RA sites and five out of the

nine RCB sites (Table 10).

Table 10. Results of logistic regression for individual aquarium fish species. For more
specific information see Appendix 2.

Fish Species	Percent of Species Found in RA Habitat Type		Percent of Species Found in RCB Habitat Type	
	N (absent)	Y (present)	N (absent)	Y (present)
Acanthurus triostegus	57.1	42.8	55.5	44.4
Ctenochaetus strigosus	14.2	85.7	0	100
Scarus dubius	14.2	85.7	0	100
Scarus psittacus	0	100	11.1	88.8
Zebrasoma flavescens	0	100	0	100
Chaetodon ornatissimus	0	100	0	100
Chaetodon quadrimaculatus	57.1	42.8	0	100
Forcipiger spp.	0	100	0	100
Acanthurus achilles	0	100	0	100
Naso lituratus	0	100	0	100
Paracirrhites arcatus	0	100	0	100
Paracirrhites foresteri	0	100	0	100
Acanthurus nigrofuscus	0	100	0	100

These results show that there is little habitat utilization preference between the RA and

RCB habitat types for these selected species. A slight preference for the RCB habitat

type is observed, but further studies are needed to verify these results.

To determine if certain substrate types were significantly more abundant in the

RA verses the RCB habitat types, the mean percentages of each small-scale substrate

type were compared using a two-sample t-test (Table 11).

Table 11. Mean percent cover for the WHAP substrate categorized by RA and RCB habitat types. * Indicates significance (p<.05).

Substrate type	RA Mean % Cover (SD)	RCB Mean % Cover (SD)	P
Boulder	0.29 (0.59)	2.26 (3.92)	0.153
Porites compressa	34.9 (13.6)	12.2 (11.1)	0.004*
Porites lobata	23.3 (9.97)	33.1 (8.96)	0.059
Pocillopora meandrina	0.17 (0.44)	0.88 (1.87)	0.277
Rubble	6.73 (6.23)	7.14 (6.68)	0.897
Sand	0.92 (1.77)	7.00 (10.9)	0.117
Dead Coral	28.6 (10.1)	32.3 (17.6)	0.601

The mean percent cover of the corals *Porites compressa* (PC) and *Porites lobata* (PL) found in each habitat type were plotted to better examine the data (See Figures 21 and 22).
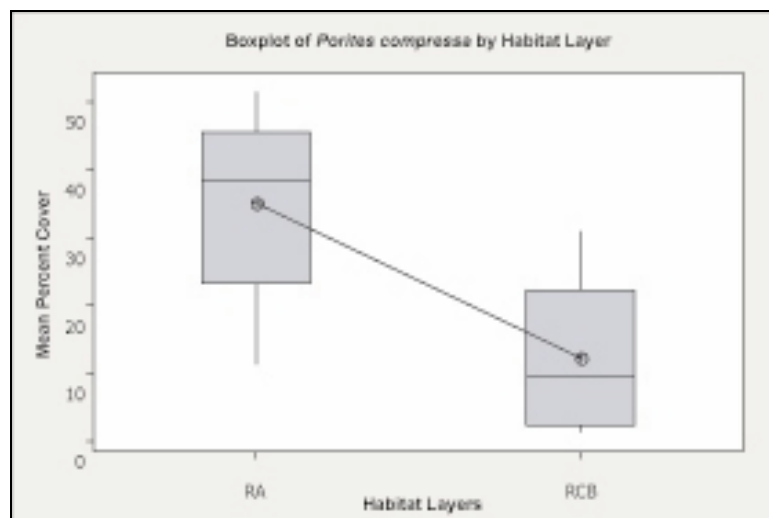

Figure 21. Boxplot illustrating the variation in the percent cover of the coral *Porites compressa* relative to the two habitat layers.
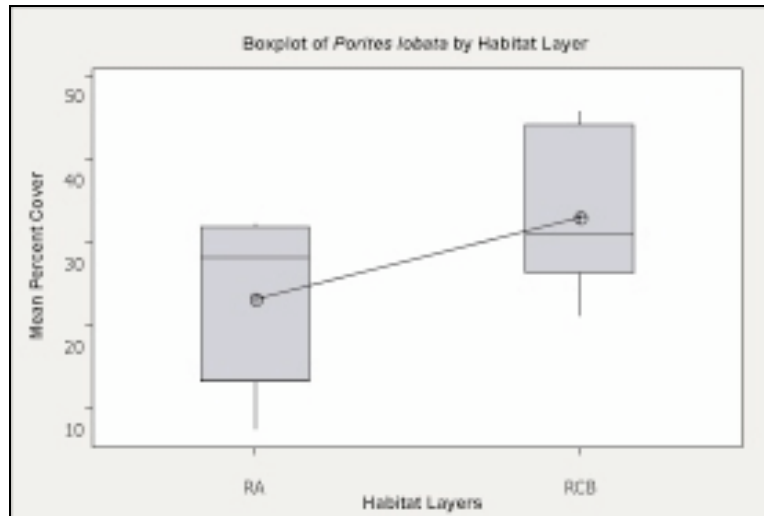
Figure 22. Boxplot illustrating the variation in the percent cover of the coral *Porites Lobata* relative to the two habitat layers.

These results show that *Porites compressa* (finger coral) was significantly more abundant (p<0.05) in the RA relative to the RCB habitat types. In addition, the *Porites lobata* substrate, while not significant (p>0.05), shows a strong trend toward being found more often in the RCB habitat type.

An additional analysis was conducted to determine if depth played a role in the large-scale habitat classification. The depth values of each site were compared to the classification of the site habitat as either RA or RBC using a two-sample t-test. The mean depth value of the RA habitat type is 41.71 $\pm$ 5.77m, and the mean depth value of the RCB habitat type is 37 $\pm$ 3.65m. These mean values were plotted to better examine the data (Figure 23).
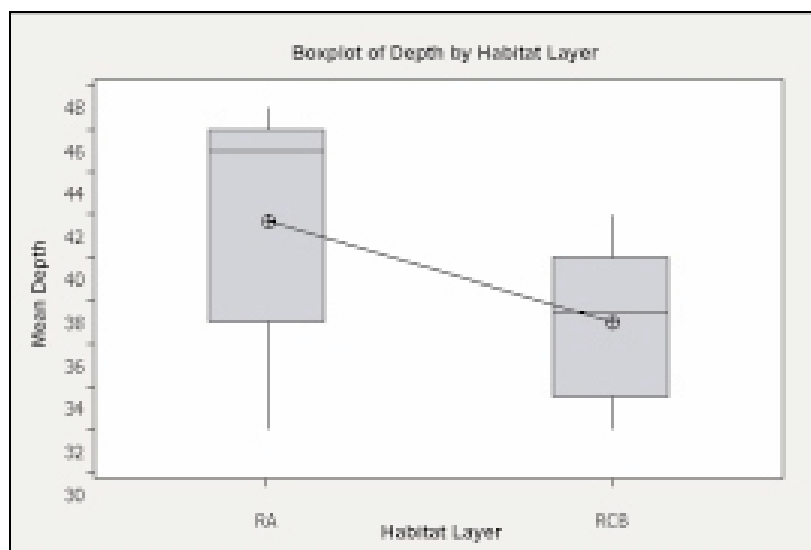
Figure 23. Boxplot illustrating how habitat layers vary with depth derived from WHAP survey data.

These results show that while depth values are not significantly different (p=0.088) between RA or RCB habitat types, the trend indicated that RA habitats occurs deeper than RCB habitat types.

Conclusions

By using the ArcGIS MDM for the analysis of habitat utilization patterns of

reef-fish, it was determined that 3 out of 5 goals were met (Table 12), and the

remaining two goals identified how the model could be modified to better meet the

needs of the marine community.

Table 12. Evaluation of goals for the MDM for this study.

Goals	Results
Goal 1: Provide a common structure for assembling, managing, and querying marine data in ArcGIS	Met
Goal 2: Provide a template for marine users	Met
Goal 3: Advance the data modeling capabilities using object-orientation	Partially met –Not met for the average user, but met for the GIS developer
Goal 4: Improve spatial data representation	Partially met—Using the MDM does not improve spatial data representation beyond what ArcMap and ArcScene already provide
Goal 5: Increase the understanding of ArcGIS	Met

To modify the MDM to better meet the needs of the marine community, it is

suggested to establish some behavioral and validation rules that will already be

programmed into the MDM. This will allow the marine community to better take

advantage of the object orientation capabilities, which would greatly increase the

analysis and modeling capabilities. In addition, it is also the suggestion of the author

to create a tool that will somehow connect to a true 3D visualization program like Java

3D or Fledermaus, and the MDM. The reasoning behind this is because the marine

data model does not improve the 2.5D in ArcScene spatial visualization of data beyond what ArcScene is already set up to do. Yet, the true 3D visualization of data is such an important component to marine work!   This can be seen in prior marine case studies where the researchers turned to the true 3D visualization program, Fledermaus, to solve the problem of visualizing very large and complex grid data (Wright, 2002), even when using ArcScene was a less expensive option. So to combine the analytical capabilities of the GIS with a true 3D visualization program would be a very useful package and perhaps provide another reason for people to use the MDM.   However, it has been argued that that this is an expensive development, that would take a lot of resources, and it may not provide a good return on the investment (Michael Blongewicz, pers. comm.).   In view of this, the author recommends using ArcGIS 9.0, which will arguably improve that spatial representation of data one step above 8.3 when using ArcGlobe, but will not yet represent the data in true 3D.

This analysis has shown that the MDM is a versatile tool that can be used to perform a variety of functions for marine users. Not only can it store, assemble and query data, but it serves as a tool for examining spatial patterns of marine data at both a large and small-scale. However, this analysis has also shown that the MDM is best suited for two types of users. The first user is one who has never used a geodatabase before. For these users, the model provides a template with complete instructions in the form of a tutorial, on how to propagate the MDM with data. This saves the time it would take to learn how to design a geodatabase from scratch. And the second type of user has prior experience with GIS and can take advantage of object-orientation. This

is because much of the analytical power behind the MDM is a result of the object-orientation functionality, which takes prior programming knowledge to fully utilize.

Ideally, the users would have a variety of marine data types to store in the geodatabase, which means that more than one feature class would be utilized. This ensures that it will be worth their time to use and personalize the model to suit their needs. On the other hand, the user would not have a great quantity of data, as the speed of queries depends on the amount of data stored in the geodatabase tables. For example, one of the tables in this study contained over 250,000 data entries. To query through the table it took the MDM over two hours.

When studying the regional habitat patterns for selected reef fish, the logistic regression showed a slight habitat preference between the RA and RCB habitat types. This is likely due to the fact that NOAA's habitat layer is very generalized, as the habitat classification shows only continuous habitats greater than 1 acre in size (Battista, 2003). The lack of finer habitat information points to a need for further research into providing more detailed habitat information at a scale between NOAA's benthic habitat delineation and WHAP's detailed substrate information. If this was available at the time of this study, it is speculated that the results for the habitat utilization pattern might have been different.

This study also found that RA habitat is highly likely to be found at deeper depths, while RCB tends to be found at more shallow depths. Additionally, the coral *Porites compressa* was found significantly more often in RA, while *Porites lobata* was found more often in RCB. This means that *Porites compressa* will more likely be found at deeper depths than *Porites lobata*. This result is consistent with other data

analyzes that have found finger coral to dominate most areas of the West Hawaii coast at 10-18m depths except along exposed headlands and recent lava flows (Dollar, 1982 and Tissot et al., 2004). This is important as finger coral provides an important habitat for juvenile aquarium fish, especially the Yellow Tang (*Zebrasoma flavescens*) (Tissot et al., 2004).

Results of the habitat analysis also showed that correlations existed between habitat information at two different scales. This result has important management implications as most marine management units are on the scale of an island or an entire state, whereas most ecological studies are typically conducted at a small-scale. Currently, the lack of detailed habitat information is a result of the high cost and time it takes to conduct a detailed survey. Thus, this study is important not only because it bridges the scale gap but it provides a new level of information to the large-scale habitat layer without high cost or time consuming surveys.

This study has also shown that depth is an important factor in habitat location. From a management prospective, this is important when looking at the boundaries of the protected areas and examining how far they extend into deeper waters, as previous studies have found that the more effective protected areas for specific species have boundaries that extend into deeper water (Tissot et al., 2004). For example, as *Porites compressa* was found more often in the RA habitat, and this coral is known to be a critical habitat for the Yellow Tang, it is important for managers to evaluate the protected area boundaries to ensure that RA habitat is also included in the protected areas.

As the MDM was successfully used as a tool to provide spatial information on the location of Hawaiian near-shore reef fish, it will be given to the HNHP to be used in conducting the MGAP. As a result, this database will be built on and used in the future. In addition, this research has identified areas of future research such as the need for more detailed habitat information, and the need to establish a stronger correlation between large-scale substrates and habitat utilization of specific fish species. Such information would be used by marine managers to evaluate the current marine protected area network and potentially establish new MPAs.

Literature Cited

Battista, Timothy. 2003. *Benthic Habitats of the Main Hawaiian Islands*. NOAA National Ocean Service: Silver Spring, MD. Preface, pgs 10-15.

Brown, Eric. 1998. Coral Reef Network: Saving Maui's Reefs, http://cramp.wcc.hawaii.edu/Study_Sites/Maui/Saving_Maui_Reefs/#contents 4/14/04.

Brown, E., E. Cox, P. Jokiel, K. Rodgers, W. Smith, B. Tissot, S.L. Coles, and J. Hultquist. 2004. Development of benthic sampling methods for the Coral Reef Assessment and Monitoring Program (CRAMP) in Hawaii. *Pacific Science* 58: 145-158.

Carr, M., J. Neigel, J. Estes, S. Andelman, R. Warner, and J. Largier. 2003. Comparing Marine and Terrestrial Ecosystems: Implications for the Design of Coastal Marine Reserves. *Ecological Applications*, 13(1) Supplement: pp S90-S107.

Christensen, J.D., C.F.G Jeffery, C. Caldow, M.E. Monaco, M.S. Kendall, and R.S. Appeldorn. 2003. Cross Shelf Habitat Utilization Patterns of Reef Fishes in Southwestern Puerto Rico. *Gulf and Caribbean Research,* 4(2): 9-27.

Dollar, S. J. 1982. Wave stress and coral community structure in Hawaii. *Coral Reefs,* 1:71-81.

DBEDT. Department of Business, Economic Development and Tourism: Statewide Hawaii GIS data, http://www.state.hi.us/dbedt/gis/organiz.htm 5/24/04, 2004.

ESRI. 2004. Downloads for Data Models, http://support.esri.com/datamodels, 4/3/04.

Friedlander, A., and E. Brown. 2003. Fish Habitat Utilization Patterns and Evaluation of the Efficacy of Marine Protected Areas in Hawaii: Integration of NOS Digital Benthic Habitat Maps and Reef Fish Monitoring Studies. NOAA Technical Report.

Hallacher, L., and B. Tissot. 1999. Quantitative Underwater Ecological Survey techniques: A coral reef monitoring workshop. Chapter 13 in: Maragos, J. E. and R. Grober-Dunsmore (eds.). Proceedings of the Hawai'i Coral Reef Monitoring Workshop, Dept. of Land and Natural Resources, Honolulu, HI.

Kleiner A., L. Gee and B. Anderson. 2000. Synergistic Combination of Technologies. Proceedings of Oceans 2000. Providence, Rhode Island: Marine Technology Society.

Leslie, H., M. Ruckelshaus, I. Ball, S. Andelman, and H. Possingham. 2003. Using Siting Algorithms in the Design of Marine Networks. *Ecological Applications*, 13(1) Supplement, pp S185-S198.

Mumby, P., J. Chisholm, C. Clark, J. Hedley, and J. Jaubert. 2001. A Bird's-eye View of the Health of Coral Reefs. *Nature*, 413:36.

NOAA 1999a. Benthic Habitats of the Main Hawaiian Islands. http://biogeo.nos.noaa.gov/products/hawaii_cd/htm/overview.htm, 4/3/04.

NOAA 1999b. Benthic Habitats of the Main Hawaiian Islands: Project Methods. http://biogeo.nos.noaa.gov/products/hawaii_cd/htm/manual.htm, 4/3/04.

PISCO, 2002. *The Science of Marine Reserves*. http://www.piscoweb.org/outreach/pubs/reserves/, 5/8/04.

Puniwai, N., S. McElvaney, and S. Hochart. 2003. *Marine Gap Analysis Project of Hawaii—Year One Final Report.* State of Hawaii Division of Aquatic Resources.

Scott, M., B. Csuti, J. Jacobi and J. Estes. 1987. Species Richness: a geographic approach to protecting future biological diversity. *BioScience.* 37(11):782-788.

SHOALS. 2000. US Army of Engineers, http://shoals.sam.usace.army.mil/Hawaii/pages/Hawaii_Big_Island.htm, 4/27/04.

Tissot, B., W. Walsh and L. Hallacher. 2004. Evaluating the Effectiveness of a Marine Reserve in West Hawaii to Improve Management of the Aquarium Fishery. *Pacific Science*, 58(2): 175-188.

Wright, D. J. 2002. *Undersea with GIS.* ESRI Press: Redlands, CA.

Wright, D.J. and Bartlett, D.J. 2000. *Marine and Coastal Geographical Information Systems*, London: Taylor & Francis, 322 pp.

Wright, D.J. and Blongewicz, M., 2003. Marine Data Model Marine Feature Classes Document, http://dusk.geo.orst.edu/djl/arcgis/diag.html, 5/17/04.

Wright, D.J., Halpin, P.N., Breman, J., and Grise, S., 2001. ArcGIS Marine Data Model Conceptual Framework. http://dusk.geo.orst.edu/djl/arcgis/frame.html, 4/20/04.

Zeiler, M. 1999. *Modeling our World*. ESRI Press : Redlands, CA

APPENDICES

Appendix 1. Marine Data Model Tutorial

Using the Marine Data Model in ArcGIS 8.3

Links for this Tutorial

Marine Data Model Home Page
Marine Data Model at ESRI
EPIC Web Data Browser
Samoa Source Data

Basics:

Introduction to the ArcGIS Marine Data Model

Tutorial Objectives

Computer and Data Requirements

Setting Up the Geodatabase:

Downloading the MDM Geodatabase

Downloading the Data

Applying the MDM Model Schema to the Geodatabase

Loading Data into the MDM Geodatabase:

Assessing Your Data and Determining

Your Database Set Up

Personalizing the MDM to Fit Your Data

Adding Data to the MDM

Adding the Geodatabase Features to your ArcMap Project:

How to add your data

Basics

Introduction to the ArcGIS Marine Data Model

The ArcGIS Marine Data Model (MDM) is a geodatabase model. It is specifically tailored to meet the needs of the marine GIS user community, by providing a standardized template into which various types of marine data can be loaded, stored, manipulated, queried and spatially analyzed. Furthermore, the geodatabase model supports object-orientation, to allow for better spatial modeling capabilities. Through object-orientation, data can become "smarter," as it can be endowed with behavioral and validation rules.

The MDM geodatabase model was generated in a series of steps, beginning with the definition of feature datasets, classes, attributes, and relationships in a Unified Modeling Language (UML) diagram created in Visio 2000. The UML diagram was then converted into a Microsoft repository format, which is an equivalent tabular structure, or schema, so that it could be loaded into Microsoft Access or other relational data servers. The schema was then applied to a personal geodatabase to create the sets of classes and attributes that were defined originally in the UML.

The geodatabase is a store of geographic data organized into a collection of datasets, feature classes, object classes and relationship classes. A feature dataset is a collection of feature classes that share a common spatial reference, such as "Marine Features." A feature class can contain point, line, or polygon features, such as "SurveyPoint" or "FeatureArea." Data tables without spatial reference can also be located in a geodatabase, such as "SurveyInfo." And finally, a relationship class is a collection of relationships between two feature classes, or a feature class and a table, such as "SurveyHasType." To make use of the newly created MDM geodatabase, data must be loaded into the appropriate feature classes and

tables. What follows is a tutorial on how to load your data into the Marine Data Model, as well as some trouble-shooting tips.

> *For a list of the MDM feature classes, go to http://dusk.geo.orst.edu/djl/arcgis/frame.html, and download the Marine Feature Classes document. Look at the "MDM classes-Sept."*
>
> *For a list of common marine data types used in the MDM, go to http://dusk.geo.orst.edu/djl/arcgis/images/marinedatatypes_final.jpg*

Tutorial Objectives

To **apply** the Marine Data Model schema to a personal geodatabase
To **prepare** your marine data for inclusion in the Marine Data Model.
To **load** your data into the Geodatabase

Computer and Data Requirements

Computer and Software:

1) ArcGIS 8.3 (untested on earlier versions)

Necessary Data Files (see instructions below for downloading this data):

Schema:
   o ArcGISMarineReposit.mdb
Instrument Data:
   o Samoa shoreline data: Tu_Coast02.shp
   o Spatial bottle data: XYBottle_Spatial.shp
   o Non-spatial bottle data: XYBottle_Info.dbf

TUTORIAL

Setting up the Geodatabase

Downloading the MDM Geodatabase

1. Open ArcCatalog.

2. Create a new folder called Samoa for your geodatabase.

3. Go to http://dusk.geo.orst.edu/djl/arcgis/diag.html and unzip the "UML Diagram and Repository." This is the schema for the marine data model. Extract it to the Samoa folder.

4. In ArcCatalog, a personal geodatabase called ArcGISMarineReposit.mbd should appear in that folder.

   Note: If you view this geodatabase outside of ArcCatalog, it should appear as a MS Access Database

Downloading the Data

1. Download the data

   For the Samoa shoreline data, go to http://dusk.geo.orst.edu/djl/samoa/

   o Download the 'Tutuila Coastline, UTM' shapefile into the Samoa folder.

   For the XYBottle Data.

   o Download and unzip the file XYBottle.zip. Make sure you have the files XYBottle_Spatial.shp and XYBottle_Info dbf.

   Note: This data was obtained from the EPIC website and has been modified for this lab.

Applying the MDM Model Schema to the Geodatabase

In order to load the ArcGIS Marine data model, the MDM schema will be applied to a personal geodatabase.

1. In ArcCatalog, click on the ArcGISMarineReposit geodatabase and

    click on the schema wizard button.

    Note: If you do not have a schema wizard button in the ArcCatalog view frame, follow the steps below:

      Go to *Tools>Customize*.

      Click the *Commands* tab.

      Scroll down to 'Case Tools' and highlight the schema wizard that appears on the right, then click close.

      What to do if the Schema Wizard does not appear in the tool commands.

      ▪ If 'Case Tools' does not appear in the list, click 'Add from file…' and browse to the Bin directory where ArcGIS was installed (usually found in c:/arcgis/arcexe83/bin).

      ▪ Add the Schema.Wiz.dll file.

      ▪ If you don't see the SchemaWiz.dll in /arcexe/bin, it may still be there but not visible. Use Tools/Find File in Windows Explorer to locate the file, and then register the .dll using RegCat.exe, which is also located in /arcexe83/bin (This too may also be invisible, follow same steps to locate it). Create a shortcut to the RegCat.exe on your desktop. Drag the SchemaWiz.dll file onto the RegCat.exe shortcut and you'll be prompted with a dialog to define where to register the .dll. Select ArcMap, ArcCatalog and ArcTools. Now, when you go to the Categories list you will see that the Case Tools option is available and the Schema Wizard icon is visible.

2. When the Schema Wizard opens, read through the first page, and hit next.

3. Click on the "Model Stored in Repository Database" and then browse to the personal ArcGIS MDM geodatabase in the Samoa folder. Click next.

4. Accept the defaults and click next again.

5. On the next page make sure that you give the geodatabase a spatial reference. This must be done here and cannot be modified later. Click on the first feature dataset "marine features." Then click on properties.

> *Tip: In deciding which coordinate system to use and how large to set the spatial extent, I find it is best to use the import option and use a shapefile/coverage. This ensures that the coordinate system is correct and the spatial extent will be large enough to accommodate your data. Knowing this be sure to choose a shapefile/coverage that covers the greatest spatial extent you will be mapping (ie If mapping the 8 main Hawaiian Islands, don't choose a shapefile of just Oahu, choose one that covers all 8 main Hawaii Islands). I stress this because once the coordinate system and spatial extent are set, they cannot be changed!*
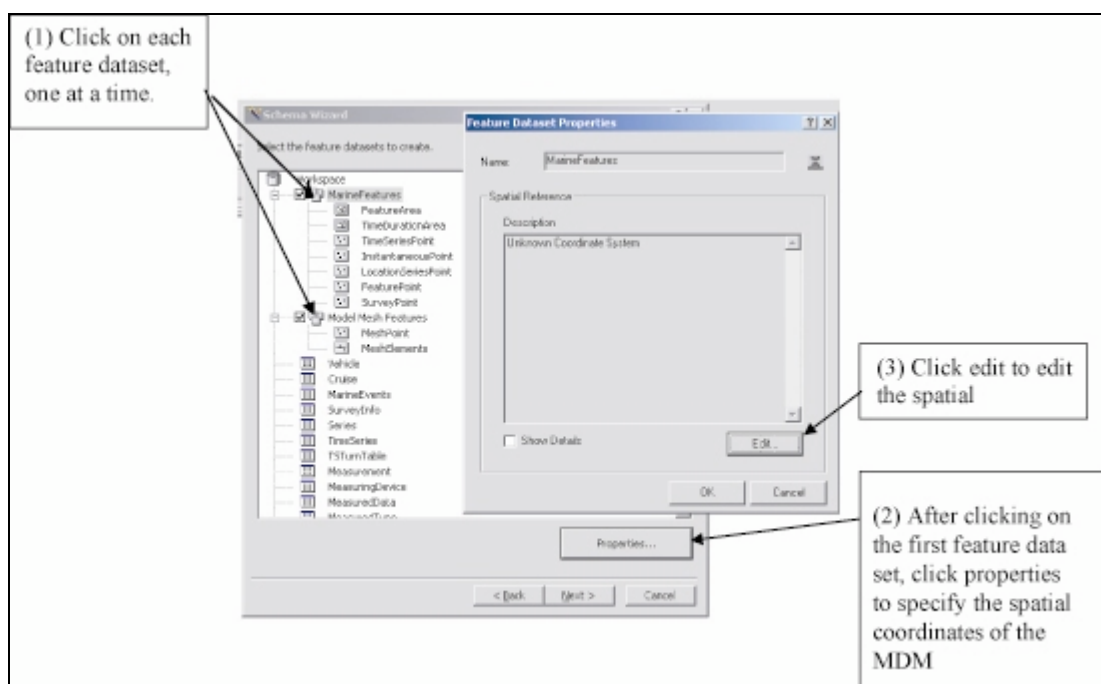
6. In the Spatial Reference Dialog, click import.

7. Navigate to the saved XYBottle_Spatial.shp.

8. Click on it to import the NAD 1983 HARN UTM Zone 2S spatial reference

9. Click next through the rest of the steps, and choose to look at the log if you want.

10. Your personal geodatabase is now ready for data.

Loading Data into the MDM Geodatabase

Assess Your Data and Determining Your Database Set Up

This is really the most important, as well as most time consuming step in working with the marine data model. However, planning the database set up now will save you much grief and stress later.

In determining how your database is set up, think about what information you have, what types of queries do you want to perform, and what can data be used as a key index to relate spatial information to non-spatial information. For this tutorial, the basic database set up is below.
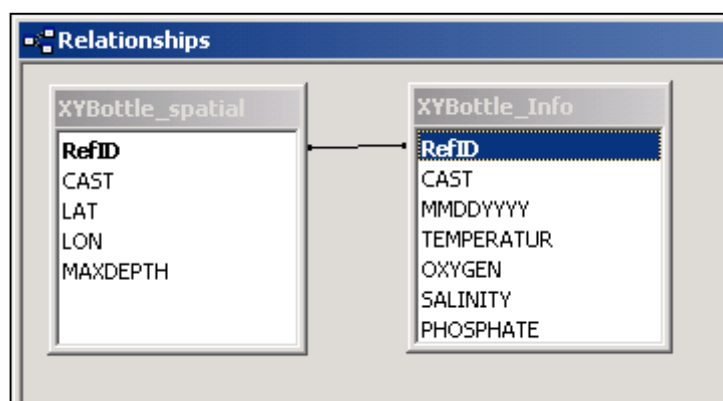
This is the database set-up, as viewed in MS Access.

This is relevant information about each shapefile that will help you later. Note the data types and fields of each shapefile. This is what you should be thinking about when applying the MDM to your own data.

Tutulia Shoreline Shapefile

> *Has spatial info*

> *Provides spatial reference*

> *Can be used in queries*

XYBottle_Spatial Shapefile

> *RefID (Long Integer)*

> *Cast (Text, 50 Characters)*

> *Lat (Double)*

> *Long (Double)*

> *MaxDepth (Double)*

The RefID will be used to relate XYBottle_Spatial data with XYBottle info (non-spatial data) in a 1-to-1 relationship, as see in the diagram above.

One reason knowing this information now is important is that to establish a relationship between two fields, they must be of long integer data type. So in this case, it is ok. However, if they were not long integers, they would need to be changed.

This database set up will allow for querying between the two tables (ie where is the temperature the greatest)

XYBottle_Info

RefID (Long Integer)

Cast (Text, 50 Characters)

MMDDYYYY (Double)

Temperature (Double)

Salinity (Double)

Phosphate (Double)

Tutulia shoreline data

1. *Examine the Tutulia shoreline shapefile in ArcCatalog. Decide which feature class it would best fit into.*

   > *Tip: Remember that point data can only be imported into point feature classes, and so on with line and polygon data.*

   > *For help go to: http://dusk.geo.orst.edu/djl/arcgis/frame.html and download the "Marine Feature Classes Document." Look at the "MDM classes-Sept." This document will give you a good idea of what data types typically would by imported into each feature class. Or take a look at the marine data model poster.*

2. *Examine the spatial coordinates of the shoreline data in ArcCatalog. Is the Tutulia shoreline in the NAD 1983 HARN UTM Zone 2S projection? As the geodatabase is in the NAD 1983 HARN UTM Zone 2S projection, all data imported into the marine data model needs to be in this projection.*

XY Bottle Data

1. Examine the XYBottle_Spatial shapefile. Determine which feature class it would be fit into.

   > Tip: In examining which feature class this data show go into, make sure to also look at the related tables, as you will likely import the XYBottle_Info dbf into a related table.

2. Examine the attributes of your data. Note the fields and the data types you have (ie RefID, Cast, latitude, longitude, maxdepth). This will become important in the next step.

3. Examine the attributes of the XYBottle_info dbf as well.

Personalizing the MDM to Fit Your Data

Once you have decided which feature classes to use (I would suggest the FeatureArea for the shoreline and the InstantaneousPoint for the spatial XYBottle data), you need to personalize the model to fit your data.

1. In ArcCatalog, right click on the FeatureArea feature class and go to properties.
2. Click on the fields tab. Here you should see the field names and data types already entered into the MDM.
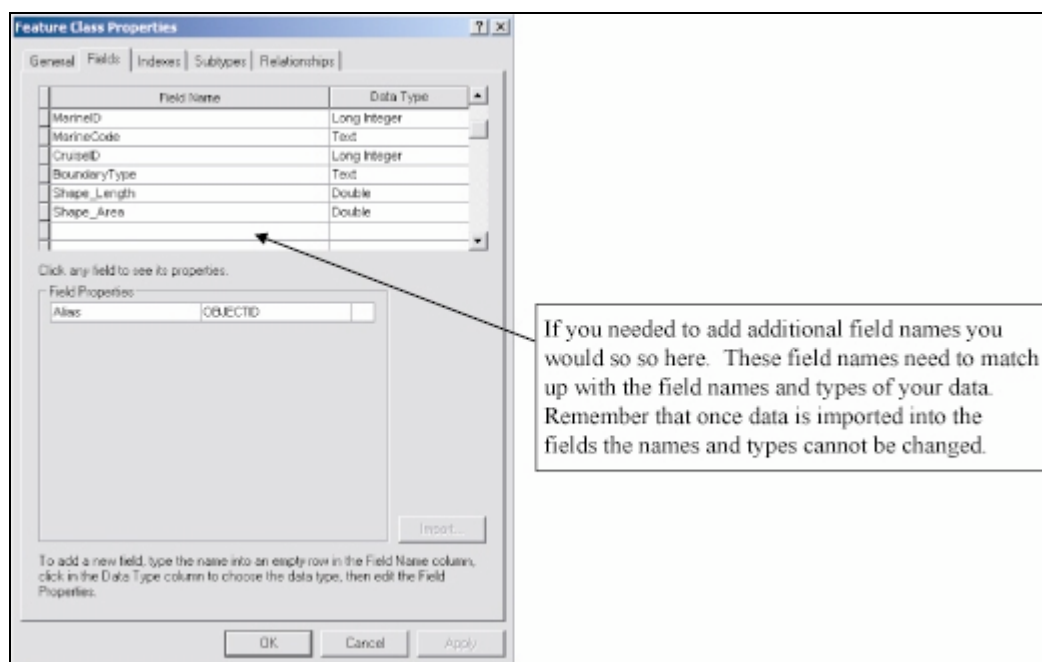
3. As the field names in the FeatureArea feature class already have an object ID and shape, this feature class does not need to be personalized before importing the Tutulia Coastline shapefile.

4. However, the InstantaneousPoint feature class does need to be personalized before importing the XYBottle data.
   Note: You data will not be imported unless you have added field names and data types to match your data.
5. To do this, follow the steps below.
   1. In ArcCatalog, right click on the InstantaneousPoint feature class and go to properties.
   2. Click on the fields tab. Here you should see the field names and data types already entered into the MDM.
   3. Add additional field names to match those in the XYBottle_Spatial dbf
   4. Add the data types (i.e. Long Integer, Double).

Tip: To determine the data types for your data, open the dbf file in access, and click on design view

Design

5. So to get you going, type in RefID in the field name column and type in Long Integer for the data type.
6. Fill in the rest of the data types for the XY bottle data.

Note: You will likely be prompted to change both cast and long field names to cast_ and long_, this is ok.

6. Your Feature Class Properties for the InstantaneousPoint should look like this.

Feature Class Properties

General | Fields | Indexes | Subtypes | Relationships

Field Name	Data Type
CruiseID	Long Integer
RecordedTime	Date
RefID	Long Integer
Cast_	Text
Lat	Double
Long_	Double
MaxDepth	Double

Click any field to see its properties.

Field Properties

Alias	Shape
Allow NULL values	Yes
Geometry Type	Point
Avg Num Points	0
Grid 1	1000
Grid 2	0
Grid 3	0
Contains Z values	No
Contains M values	No
Default Shape field	Yes
Spatial Reference	NAD_1983_HARN_UTM_

Import...

To add a new field, type the name into an empty row in the Field Name column, click in the Data Type column to choose the data type, then edit the Field Properties.

OK | Cancel | Apply

7. Accept the rest of the defaults by clicking next through all the prompts.

Modifying the Measurement Table
As the Instantaneous Point feature class has a pre-established relationship with the measurement table, this is table you are going to load the XYBottle_info (or the non-spatial) data into.

1. Follow the same steps as above to personalize this table in the geodatabase, except right click on the "Measurement" table.

Adding Data to the MDM

NOTE: For instructions on how to get your own personal data from a .dbf file with decimal degree coordinates to into a feature class, go to the end of the tutorial.

Adding the Tutulia shapefile
1. In ArcCatalog, right click on the FeatureArea feature class, go to load data.
2. Navigate to the Tutulia shapefile, click add, then next.
3. Accept the defaults on the first page and click next.
4. Accept the rest of the defaults by clicking next.
5. Double check that the data added successfully, by viewing the shapefile in ArcCatalog.
   Tip: It may take a minute to update the newly imported info into the geodatabase. If the Tutulia shapefile does not appear right away, go to view>refresh, and then view your Tutulia shapefile.

Loading the XYBottle_Spatial.shp
1. Right click on the InstantaneousPoint feature
class and go to load data.
2. Navigate to the XYBottle_Spatial.shp, and click
next two times.
3. Now match up your data.
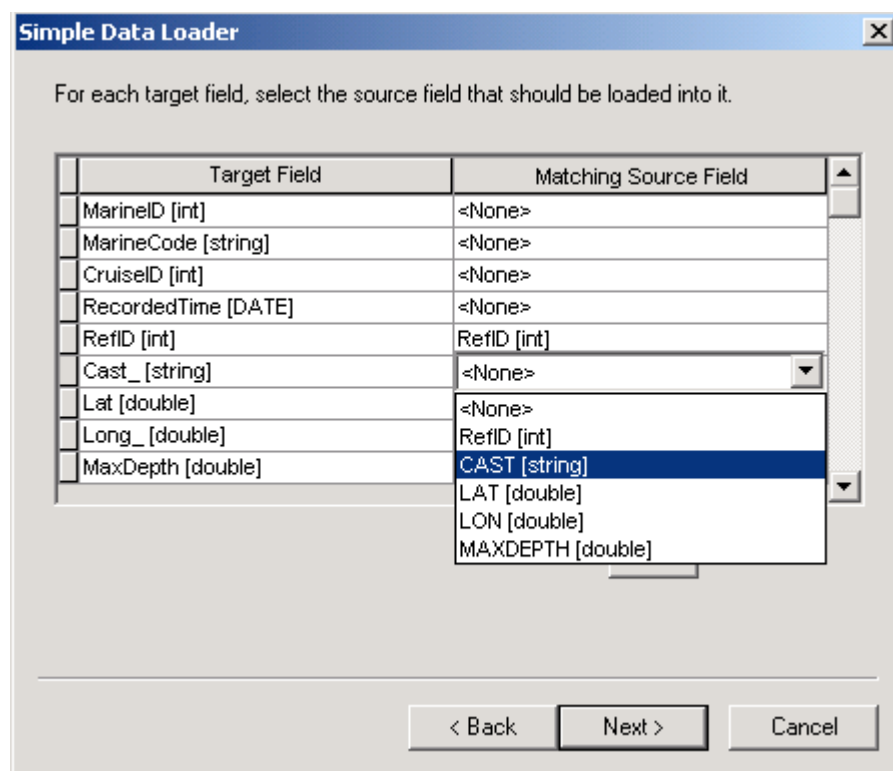
4. Click next through the last screens and then finish. Now your data is
loaded into the feature class.

Adding the XYBottle_Info
   1. To add the XYBottle_Info to the Measurement table, follow the
   same 4 steps above, only right clicking on the Measurement table.

Creating a Relationship Class
   While there is a predefined relationship class between the
      InstantaeousPoint feature class and the Measurement table, this
      relationship is a 1-many. This means that you will need to create
      your own 1-1 relationship class.
   1. Right click on the ArcGISMarineDataReposit.mbd, go to
      new>relationship class.
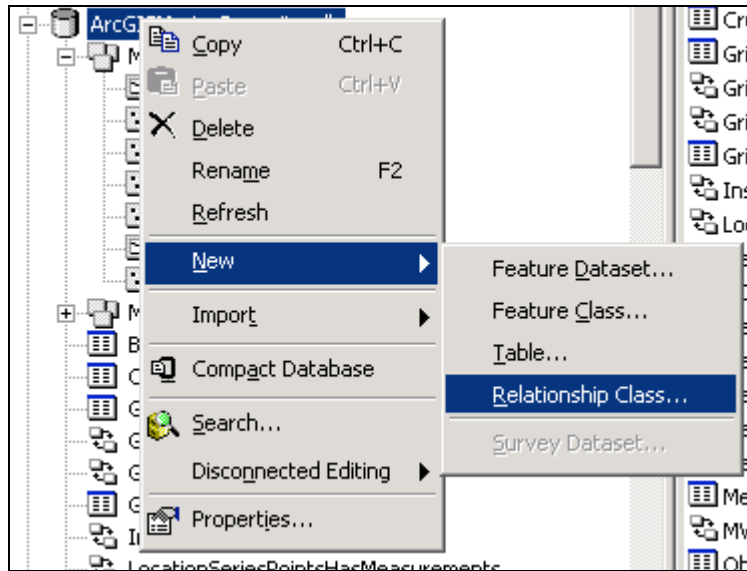
2. Name your new relationship, InstantPT_Measure.
3. The origin table is the InstantaneousPoint and the destination table
   is the Measurement table.
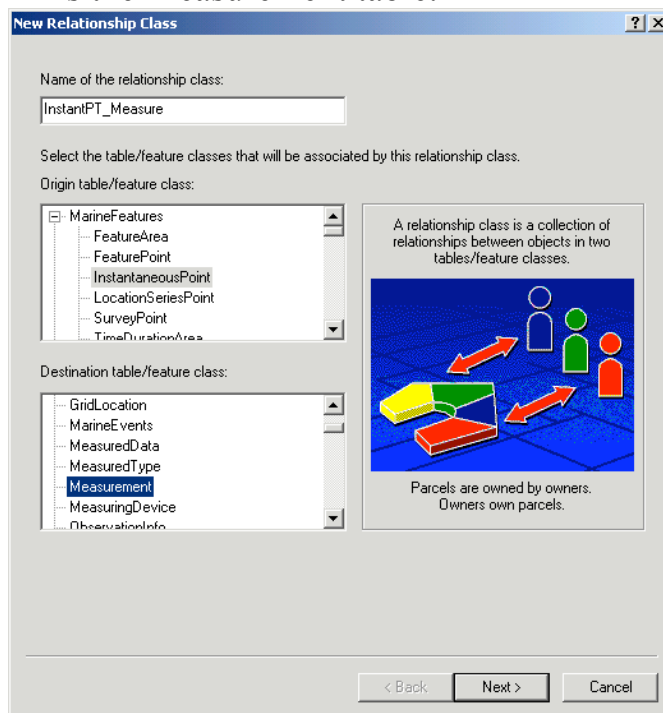
4. Click next through the screens, accepting the defaults but making
sure it is a:

        simple relationship
        origin/destination tables are right
        relationship

no attributes added to the relationship

5. On the next screen, select refID as this serves to relate your two tables.

6. Click next, then finish.

5. Once the InstantPT_Measure relationship class appears in ArcCatalog the two items should be related.

>   To check that relationships are correct, use the info button in ArcMap. The relationships between data should appear in the info pop-up. Otherwise, open up the feature class attribute table and check relationships by going to *Options>related tables*.
>   If the relationships are not seen, try re-opening your ArcMap project again, as ArcMap does not always refresh very quickly.

Adding the Geodatabase Features to your ArcMap Project

How to add data to your ArcMap Project

1. Add the two feature classes, FeatureArea and InstantaneousPoint, from the geodatabase

2. Add the Measurement table.

3. Now you are ready to query for information through the tables.

4. Just to practice, find out how many points have a temperature of 26 degrees.

   To do this, open the Measurement table. Highlight the columns that contain temperature data of 26 degrees.

    In the attribute table, go to options>related tables>XYBottle_Spatial

    This will highlight all the points that have a temperature of 26 degrees.

5. You are done. Congrats.

Follow these steps to go from a .dbf file with decimal degree coordinates to importing it into a feature class

1. Open ArcMap and add the dbf.

2. Go to tools>add xy data
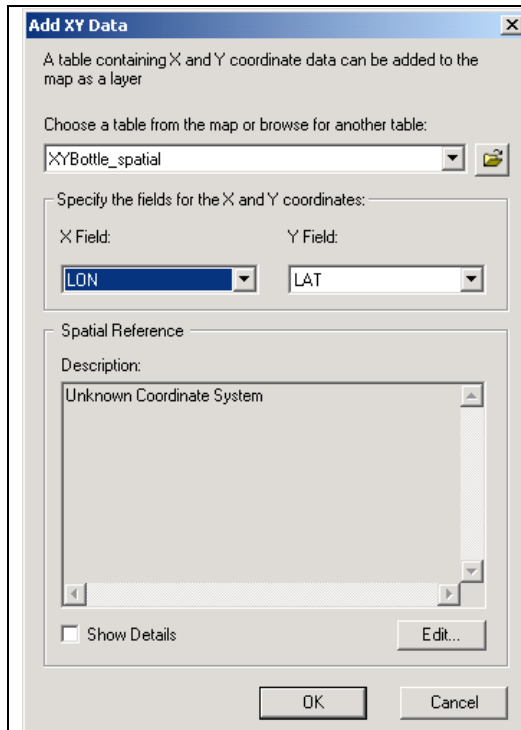
4. Set the coordinate system to match that in which your data was collected in

5. Click next through the rest of the prompts.

6. Now something called an "Events" layer that appears in your project

7. Save this as a shapefile, by right clicking and going to >data>export. Save it to a folder.

Now you will likely need to reproject this shapefile to match the coordinate system and spatial extent of your geodatabase. To do this, follow these steps:

1. Open ArcTool box

2. Go to Project Wizard, and launch the wizard.
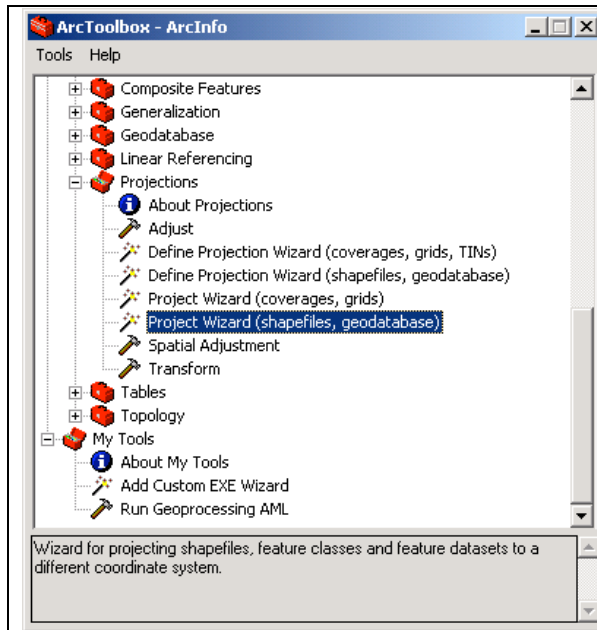
3. Navigate to your new shapefile, click next.

4. Save it to a folder, and give it a new name.

5. Give your shapefile a coordinate reference.

    Note: You might need to transform it.

8. Click next through the remaining screens, then finish.

9. Now your shapefile is ready to be added to feature classed.
    Note: Some feature classes in the MDM require that the shapefile is
    3D!

This Tutorial Created 4/8/04 by Alyssa Aaby, revised 5/29/04.
Department of Geosciences, Oregon State University

Appendix 2. Logistic Regression

The data from the logistic regression that shows the number of present or absent

species in each habitat type.

Fish Species	Present in Habitat	N	Y	Total
Acanthurus triostegus	RA	4	3	7
	%	57.14	42.86	100
	RCB	5	4	9
	%	55.56	44.44	100
	Total	9	7	16

Fish Species	Present in Habitat	N	Y	Total
Ctenochaetus strigosus	RA	1	6	7
	%	14.29	85.71	100
	RCB	0	9	9
	%	0	100	100
	Total	1	15	16

Fish Species	Present in Habitat	N	Y	Total
Scarus dubius	RA	1	6	7
	%	14.29	85.71	100
	RCB	0	9	9
	%	0	100	100
	Total	1	15	16

Fish Species	Present in Habitat	N	Y	Total
Scarus psittacus	RA	0	7	7
	%	0	100	100
	RCB	1	8	9
	%	11.11	88.89	100
	Total	1	15	16

Fish Species	Present in Habitat	N	Y	Total
Zebrasoma flavescens	RA	0	7	7
	%	0	100	100
	RCB	0	9	9
	%	0	100	100
	Total	0	16	16

Fish Species	Present in Habitat	N	Y	Total
Chaetodon ornatissimus	RA	0	7	7
	%	0	100	100
	RCB	0	9	9
	%	0	100	100
	Total	0	16	16

Fish Species	Present in Habitat	N	Y	Total
Chaetodon quadrimaculatus	RA	4	3	7
	%	57.14	42.86	100
	RCB	0	9	9
	%	0	100	100
	Total	4	12	16

Fish Species	Present in Habitat	N	Y	Total
Forcipiger spp.	RA	0	7	7
	%	0	100	100
	RCB	0	9	9
	%	0	100	100
	Total	0	16	16

Fish Species	Present in Habitat	N	Y	Total
Acanthurus achilles	RA	0	7	7
	%	0	100	100
	RCB	0	9	9
	%	0	100	100
	Total	0	16	16

Fish Species	Present in Habitat	N	Y	Total
Naso lituratus	RA	0	7	7
	%	0	100	100
	RCB	0	9	9
	%	0	100	100
	Total	0	16	16

Fish Species	Present in Habitat	N	Y	Total
Paracirrhites arcatus	RA	0	7	7
	%	0	100	100
	RCB	0	9	9
	%	0	100	100
	Total	0	16	16

Fish Species	Present in Habitat	N	Y	Total
Paracirrhites foresteri	RA	0	7	7
	%	0	100	100
	RCB	0	9	9
	%	0	100	100
	Total	0	16	16

Fish Species	Present in Habitat	N	Y	Total
Acanthurus nigrofuscus	RA	0	7	7
	%	0	100	100
	RCB	0	9	9
	%	0	100	100
	Total	0	16	16