

December 14, 2004

Dr. Francisco Dallmeier, Director  
Monitoring and Assessment of Biodiversity Program  
Smithsonian Institution  
Washington, DC 20005

Dear Dr. Dallmeier:

Enclosed is our report entitled Biodiversity Database Management. It was written at the Smithsonian Institution during the period October 25 through December 14, 2004. Preliminary work was completed in Worcester, Massachusetts, prior to our arrival in Washington DC. Copies of this report are simultaneously being submitted to Professor Demetry and Petruccelli for evaluation. Upon faculty review, the original copy of this report will be catalogued in the Gordon Library at Worcester Polytechnic Institute. We appreciate the time that you, Dr. Alfonso Alonso, Ms. Jennifer Sevin, and Mr. Patrick Campbell have devoted to us.

Sincerely,

Emmanuel Fernandez

Michael Itz

Cory Sullivan

BIODIVERSITY DATABASE MANAGEMENT

Report submitted to:

Joseph D. Petruccelli and James S. Demetry

Worcester Polytechnic Institute

Washington, Project Center

by

Michael Itz

\_\_\_\_\_

Emmanuel Fernandez

\_\_\_\_\_

Cory Sullivan

\_\_\_\_\_

In Cooperation With

Ms. Jennifer Sevin, Education and Training Director

Smithsonian Institution, National Zoological Park

Date: December 14, 2004

\_\_\_\_\_  
Professor Joseph D. Petruccelli, Advisor

\_\_\_\_\_  
Professor James S. Demetry, Co-Advisor

## **Abstract**

This report, prepared for the Smithsonian Institution's Monitoring and Assessment of Biodiversity Program (SI/MAB), describes the development and implementation of a procedure for internal vegetation plot data handling and organization. It outlines a process for continued collection, storage, and eventual distribution of the data. The report includes a comprehensive manual and associated software tools developed to facilitate the described process.

# Authorship Page

|                          |  |
|--------------------------|--|
| <b>Executive Summary</b> | Emmanuel Fernandez, Cory Sullivan  |
| <b>Introduction</b>      | Emmanuel Fernandez, Michael Itz, Cory Sullivan   |
| <b>Background</b>        | Emmanuel Fernandez, Michael Itz, Cory Sullivan   |
| <b>Methodology</b>       | Michael Itz, Cory Sullivan   |
| <b>Results</b>           | Michael Itz  |
| <b>Conclusions</b>       | Michael Itz  |
| <b>Recommendations</b>   | Cory Sullivan  |
| <b>Appendix A</b>        | Emmanuel Fernandez, Michael Itz, Cory Sullivan in coordination with Meryl Gray, Vincent Papia, and Peter Vallieres of the Biodiversity Education and Training Initiative |
| <b>Appendix B</b>        | Emmanuel Fernandez, Michael Itz, Cory Sullivan   |
| <b>Appendix C</b>        | Cory Sullivan  |
| <b>Appendix D</b>        | layout by Cory Sullivan, generated from software by Michael Itz  |
| <b>Appendix E, F</b>     | Emmanuel Fernandez, Michael Itz, Cory Sullivan with revisions by MAB Assistant Director Alfonso Alonso   |
| <b>Appendix G</b>        | Emmanuel Fernandez, Michael Itz, Cory Sullivan with revisions by MAB Education/Training Coordinator Jennifer Sevin   |
| <b>Appendix H</b>        | Michael Itz  |

## Acknowledgments

The Biodiversity Database Management project team would like to thank the following for their assistance in making this project possible:

From MAB:

- **Francisco Dallmeier**, Director
- **Alfonso Alonso**, Assistant Director
- **Patrick Campbell**, Research Ecologist
- **Jennifer Sevin**, Education and Training Coordinator

From WPI:

- **Profs. Demetry and Petruccelli**, project advisors
- The *Smithsonian Biodiversity Education and Training Initiative* project team

# Table of Contents

|   |      |
|---|------|
| Title Page .....  | i    |
| Abstract .....  | ii   |
| Authorship Page .....   | iii  |
| Acknowledgments .....   | iv   |
| List of Figures .....   | vii  |
| List of Tables .....  | viii |
| 1 Executive Summary .....                                       | 1    |
| 2 Introduction .....  | 4    |
| 3 Literature Review .....                                       | 7    |
| 3.1 Monitoring and Assessment of Biodiversity Program .....     | 7    |
| 3.1.1 Biodiversity Plots .....                                  | 7    |
| 3.1.2 Plot Methodology .....                                    | 8    |
| 3.2 Data Management .....                                       | 8    |
| 3.2.1 Database Concepts .....                                   | 9    |
| 3.2.2 Forest Inventory and Analysis .....                       | 10   |
| 3.2.3 BioMon .....  | 11   |
| 3.2.4 SALVIAS .....   | 14   |
| 3.3 Data Processing .....                                       | 18   |
| 3.3.1 Cleaning Data .....                                       | 19   |
| 3.3.2 Statistical Analyses .....                                | 21   |
| 3.3.3 Java for Manipulating Data .....                          | 23   |
| 4 Methodology .....   | 25   |
| 4.1 Improvement of Internal Access to Plot Data at MAB .....    | 25   |
| 4.1.1 Organize Data into Regions .....                          | 26   |
| 4.1.2 Test Run of Data Movement .....                           | 28   |
| 4.1.3 Move Data Formats to BioMon .....                         | 29   |
| 4.1.4 Harmonize Taxonomic Names, Authorships, and Codes .....   | 30   |
| 4.1.5 Produce Summary Data .....                                | 35   |
| 4.2 Increase the Quantity of Plot Data MAB Holds .....          | 35   |
| 4.2.1 Sites with Previously Submitted Data .....                | 36   |
| 4.2.2 Planned Sites without Submitted Data .....                | 36   |
| 4.2.3 SI Biodiversity Training Course Participants .....        | 36   |
| 4.3 Make the Data Available to Researchers Outside of MAB ..... | 37   |
| 4.3.1 Send Data Sets to SALVIAS .....                           | 37   |
| 4.3.2 Publish Summary Data .....                                | 38   |
| 4.4 Materials Produced for Future Work .....                    | 38   |
| 5 Results and Discussion .....                                  | 40   |
| 5.1 Improvements to Access at MAB .....                         | 40   |
| 5.1.1 Complete Inventory .....                                  | 40   |
| 5.1.2 Organized Regional Databases .....                        | 41   |
| 5.1.3 Procedure for Maintaining Organization .....              | 42   |
| 5.1.4 Software Tools to Aid Organization .....                  | 42   |
| 5.2 Additional Data for MAB .....                               | 44   |
| 5.2.1 Course Participant Contacts .....                         | 44   |

|  |     |
|--|-----|
| 5.2.2 Researcher Contacts.....                                 | 44  |
| 5.3 Increased External Availability.....                       | 45  |
| 5.3.1 Web Access to Summaries.....                             | 45  |
| 5.3.2 SALVIAS Access to Data.....                              | 46  |
| 6 Conclusions.....   | 47  |
| 6.1 Regarding Internal Data Management.....                    | 47  |
| 6.2 Regarding Increasing the MAB Plot Collection.....          | 48  |
| 6.3 Regarding External Data Availability.....                  | 48  |
| 7 Recommendations.....   | 50  |
| Incorporate new data sets as they are received.....            | 50  |
| Continue contacting known sources of data.....                 | 52  |
| Continue to make data available.....                           | 53  |
| 8 References.....  | 54  |
| 9 Appendices.....  | 55  |
| Appendix A: Sponsor Description.....                           | 56  |
| Appendix B: Pre-Project Plot Inventory.....                    | 61  |
| Appendix C: Data Handling Manual.....                          | 66  |
| Appendix D: Summary Analyses.....                              | 110 |
| Appendix E: Previously Sent Data Letter.....                   | 117 |
| Appendix F: Intent to Establish Biodiversity Plots Letter..... | 118 |
| Appendix G: Education Course Participants Letter.....          | 119 |
| Appendix H: Software Tools and Other Electronic Resources..... | 120 |

## List of Figures

|  |    |
|--|----|
| Figure 1: Entry of plot session parameters in BioMon .....                                 | 13 |
| Figure 2: Entry of sample data in BioMon .....   | 14 |
| Figure 3: TaxonScrubber interface .....  | 21 |
| Figure 4: First step in merging, converting all data to BioMon format.....                 | 30 |
| Figure 5: Simplified example of combining species lists .....                              | 32 |
| Figure 6: Simplified example of adjusting species codes for plots in previous figure ..... | 33 |
| Figure 7: Before and after species list organization.....                                  | 34 |
| Figure 9: Screenshot of interface for merging species lists .....                          | 43 |
| Figure A1: MAB organizational chart .....  | 59 |



## List of Tables

|  |    |
|--|----|
| Table 1: SALVIAS full-access sample description data.....                                    | 17 |
| Table 2: SALVIAS access permissions .....  | 18 |
| Table 3: First 10 rows of BioMon statistics for Luquillo National Forest.....                | 23 |
| Table 4: Except from inventory, North American plots .....                                   | 27 |
| Table 5: Distribution of MAB censuses by continent.....                                      | 28 |
| Table 6: Distribution of MAB censuses by region and format.....                              | 29 |
| Table 7: Distribution of MAB censuses by data format .....                                   | 29 |
| Table 8: Distribution of MAB censuses by regional database.....                              | 40 |
| Table 9: Code counts for each regional database for different levels of identification ..... | 41 |

# 1 Executive Summary

The Smithsonian Institution, through the office of Monitoring and Assessment of Biodiversity (MAB), possesses a collection of vegetation plot data from locations around the world. The plot data are recorded into the Smithsonian's BioMon database suite by researchers in the field. After years of gathering data, analysis of the database became complex due to the difficulties associated with accessing large quantities of data within BioMon. The WPI project team was given the task of devising a solution which would better manage plot data held at MAB while taking into account future data submissions and allowing for dissemination through electronic means. Therefore, the goal of this project was to (1) improve internal access to vegetation plot data at MAB; (2) get access to existing vegetation plot data that can be incorporated into the MAB network; (3) make the data available to individuals outside of MAB.

To improve the data management at MAB the team first assessed the existing situation, including data formats and organization. This required an inventory to identify the plot data held on computers and media in the MAB offices. The team, then organized the data by geographic region to facilitate merging the data into a consistent BioMon structure.

Through the course of the project the data cleaning and organization steps demanded a significant amount of time. Data sets were converted to BioMon format, underwent standardization and merging of their contained species lists into regional master species lists, were imported into larger regional BioMon databases, and summarized in brief reports for ease of reference. Standardizing species lists was most problematic due to taxonomic names which were either misspelled, incorrectly identified, or unidentified altogether. Three

software tools were developed by the team to ease the species cleaning, data merging and summary production steps. The tools significantly reduced the amount of time it took to organize a data set and generate its summary documentation. All data handling steps were documented in the form of a handbook for future users. With the combination of the data handling manual, software programs, and inventory of MAB's data as legacy tools, MAB will be able to further pursue its efforts toward increasing and diffusing biodiversity information.

Concurrently, the team made efforts to identify potential sources for additional data. It identified three such sources: researchers who previously submitted plot data, researchers who have plot data noted by MAB though not stored at MAB, and participants from education courses conducted by MAB who may have gathered data after their training. The team emailed letters to the education course participants. For the other two groups, due to relationships MAB has with other research programs, MAB researchers were asked to make contact with letters similar to those prepared by the team.

The team also pursued means by which the data MAB collects could be made more widely available to outside parties. Of these, particular focus was given to providing the general public with brief overview summaries of the plot data via the web and providing other researchers with the raw data via the SALVIAS network at the University of Arizona. The process to generate summaries was outlined and supported by software, and a method to package data for transmission to SALVIAS was developed and described.

The team drew several conclusions regarding the internal handling of data at MAB, efforts to obtain existing data, and the external availability of MAB's data. First, a definition of the data organization process, the continued integration of new data sets into MAB's

database, and the assistance of software in the data handling process will complement data handling at MAB. Therefore, we recommend that MAB implement the procedure described in the data handling manual, investigate software improvements for database storage, incorporate new data sets as they are received, and track metadata within their plot databases. Second, through contacting researchers who have or potentially have data of interest to MAB, the amount of data available at MAB will increase. Therefore, we recommend that MAB maintain researcher contact information by periodically contacting researchers and MAB training course participants about additional plot data. Finally, to satisfy outside interest in MAB's plot data, we recommend that MAB post and continually update summary documents on their website and submit all existing and any future data sets to the SAVLVIAS research network.

## 2 Introduction

Current trends in global warming, atmospheric pollution, and a decline in general planetary well-being threaten the Earth and all of its inhabitants. The continued availability of renewable natural resources within threatened regions depends on the stability of ecosystems influenced by these trends. Attempts to ensure ecological stability start with studying biodiversity in those threatened regions. Data collected from ecologically sensitive areas can be used to further studies of biodiversity and how it changes over time. Looking for trends in vegetation data over time, measures which tend to be stationary yet sensitive to change, allows one to more clearly understand an ecosystem's current health. In particular, biologically diverse ecosystems in South America and Africa are under study by groups such as the Monitoring and Assessment of Biodiversity (MAB) Program at the National Zoo. Making available the information gained through those studies enables individuals and groups to take action to prevent further ecological damage. Efforts to combat ecological instability, including those by the MAB, must therefore answer questions regarding how to organize and distribute their data.

According to A. Alonso and J. Sevin at MAB (personal communication, September 9, 2004), the Smithsonian Institution possesses an extensive collection of sampled vegetation plot data from sites around the world. At the start of our project, these raw data were collected and awaiting more intensive analysis. Proper organization and subsequent availability to researchers of these biodiversity data were the logical next steps towards making that analysis possible. Ultimately, through analysis these data will yield a greater understanding of how to utilize the natural resources of their respective regions. Therefore,

MAB was preparing to revise and organize the database to make it available on the Internet where it can be utilized by researchers participating in other research networks.

MAB has gathered plot survey data since 1987. To aid in this process, MAB worked to develop a data management package, called BioMon, to facilitate collecting, cleaning up, and storing vegetation plot data. BioMon has seen two revisions, with the most recent 1992 version in use today. A significant portion of the MAB data had already been gathered in separate BioMon databases by the start of this project, though plots were collected in several other data formats as well. In addition, MAB had been in communication with the vegetation data repository SALVIAS (Synthesis and Analysis of Vegetation Inventories Across Scales) to make the data available for research pertaining to tropical areas where vegetation plots have been established.

When MAB's biodiversity plot data collection began in the 1980s, the technological potential for worldwide Internet distribution was unforeseen and thus not planned for in definitions of protocols and structures for data collection (J. Sevin and A. Alonso, personal communication, September 9, 2004). Efforts have since been made to incorporate such opportunities into MAB's procedures, including the development of the BioMon database application. Though BioMon mainly focuses on the data collection and organization steps, it can serve as a foundation for the eventual distribution of vegetation plot data. Further development of data distribution procedures requires more focused study of the organizational process used with vegetation plot data and how potential BioMon users could exchange data with other field researchers and networks such as SALVIAS.

The goal of this project was to improve internal access to plot data at MAB, obtain additional existing vegetation plot data that could be incorporated into the MAB network,

and make the data available to researchers outside of MAB through electronic channels. The first step consisted of a thorough examination of the data already present at the MAB central office. Next, the existing data were entered into the BioMon database according to geographic region to create a standardized method for data storage. Data summaries for each studied region were then produced and published electronically on the MAB web site for the general public. Meanwhile, the team contacted known potential plot data sources to further increase the volume of data in MAB's possession. Finally, the team established means to transmit MAB's plot data to SALVIAS for further distribution among researchers. The steps taken to move data from its reception by MAB, its incorporation into BioMon, and its dispatch to SALVIAS were documented for future use in a manual for MAB data handlers. The project thus yielded a streamlined process for handling data that will enable researchers, decision makers, students, and other potential recipients of the data to better take advantage of the data in their work.

## **3 Literature Review**

### ***3.1 Monitoring and Assessment of Biodiversity Program***

The Monitoring and Assessment of Biodiversity Program (see Appendix A), part of the Conservation and Research Center of the Smithsonian Institution's National Zoological Park, currently possesses one of the best collections of vegetation plot data in the world (A. Alonso, personal communication, September 9, 2004). These high-quality, scientifically-verified data are obtained periodically from a worldwide network of contributors. With these data, MAB fuels research and education programs to provide tools to “governments, industries, academia, nongovernmental organizations, local communities, and others” in pursuit of long-term “sustainable use of natural resources” within their regions (National Zoological Park, 2004a, Monitoring and Assessment of Biodiversity). The ability of the organization to continue its scientific research to feed and target its education programs depends on its continued ability to gather information about the global status of biodiversity.

#### **3.1.1 Biodiversity Plots**

To measure and otherwise track biodiversity, the MAB Program manages vegetation data regarding species composition, forest structure and dynamics (recruits and disappearance), and distribution of trees within almost 300 research plots around the world. Plots are selected for the representation of species and habitats, and surveyors and botanists then collect information about the plot according to protocols established by F. Dallmeier and J. Comiskey (1996).



### **3.1.2 Plot Methodology**

Dallmeier and Comiskey (1996) break one hectare (100m<sup>2</sup>) plots into twenty-five square quadrats. Within each quadrat a team of three researchers locate every stem of every tree of a minimum diameter at breast height (DBH). The DBH is usually four or ten centimeters taken at a standardized height of 1.3 meters. They measure the tree, mark it with a unique number, and map its location to approximately one half meter accuracy. At least five specimens are taken as vouchers for proper species identification. They are collected, brought back, dried out, and deposited in herbaria. High resolution photocopies are taken such that the leaves of identified trees can be compared to specimens in the field.

### **3.2 Data Management**

As MAB (National Zoological Park, 2004a) has recognized, “[t]he most efficient way to manage biological information is with a relational database” (Relational Databases), though few working on the data have much skill in databasing software. Technical solutions to managing large amounts of data have been the primary concern of much computer software since the dawn of modern computing, the end result of which has been the wide variety of modern database management systems. The gamut includes both stand-alone systems, such as Microsoft Access, where the user interface is directly connected to the data storage, and database servers such as MySQL or Oracle, where the database is served over a network to other client programs which provide the interface. A well-designed database can be a powerful tool for researchers. Therefore, as it concerned MAB’s current data storage--including the relational databases used by BioMon--this project benefited from some knowledge of how databases are organized and operated upon.

### **3.2.1 Database Concepts**

Before any work can be done upon the data sets, a system of organization must be selected to store them. This system will impose a structure upon the data and dictate how they are physically stored by the backing software. This data model should “capture the universe that it is representing as accurately, completely, and naturally as possible” (Thuraisingham, 1997, p. 26) so that users will be able to interact with it in an intuitive way. The two primary models used for viewing data today are the relational model and the entity-relationship-attribute model. Other models, including the hierarchal model and the object oriented model, are less often used in modern, general-purpose database management systems.

#### ***3.2.1.1 The Relational Model***

Many commercial software applications today use a relational model for capturing data. This system, as outlined by Ramakrishnan (2000, pp. 20-25), breaks the data down into a set of relations, called tables, which have rows, often called tuples. Each element within the tuples is atomic in that it is a single number or string. One may perform operations such as “select”, “join”, and “project” upon the relations to retrieve, connect, and narrow the scope of data.

#### ***3.2.1.2 The Entity-Relationship-Attribute Model***

As described by Thuraisingham (1997 pp. 30-31), the Entity-Relationship-Attribute (ERA) model adds a layer of abstraction over the simple relational database by defining entities, relationships between those entities, and attributes of those entities. Entities stand in for the real-world concepts being represented, for example trees and plots. Relationships connect entities together in meaningful ways: for example a plot contains many trees.

Relationships may be one to many (as with the trees and their plot), many to many, or simply one to one. Entities may also have attributes, such as the latitude of the plot or the species of the tree. However, as Ramakrishnan (1997, pp. 25) indicates, modern systems often do not directly support the ERA model for actual implementation purposes, requiring analysis to generate a corresponding purely relational structure and leaving the ERA model mainly useful for planning and conceptualizing structure.

### **3.2.2 Forest Inventory and Analysis**

The Forest Inventory and Analysis (FIA) Program of the USDA Forest Service performs a census of trees in sample locations throughout the United States. Foresters systematically collect information about existing trees within a given plot area every 5 to 7 years to better determine the health of both forest and non-forest regions (Forest Inventory and Analysis, 2004, Program Background). All 50 states and some U.S. territories are included in this inventory. The FIA interests include not only the measuring and protection of land resources but also the sustained utilization of such renewable resources.

The FIA collects data resembling those of interest to MAB, including species, DBH, location, and overall health. They also target an audience similar in composition to that of the MAB, including internal analysts, federal and state agencies, people in academia, forest industry consultants, and nature conservation groups. Because of these similarities, the approach of the FIA should serve as an example of an existing process for collecting and distributing ecological data.

According to Northeastern FIA Information Management Section head Tom Frieswyk (personal communication, September 28, 2004), the FIA attempts to electronically enter the data while in the field or as close to it as possible by using electronic collection

software. The data are then collected by the regional office, cleaned up, merged together, and then analyzed as a complete set for various characteristics. Finally, each regional section submits its data to a central server at the University of Minnesota which stores them and provides public access via the web.

This three-step process of collection, processing, and distribution of data sets can be generalized from the FIA and applied to MAB's operations. BioMon provides electronic field collection and a medium for transport of data. The SALVIAS Project at the University of Arizona provides means to deliver the data to a larger audience. To complete the chain, the MAB needed to define the middle processing step.

### **3.2.3 BioMon**

To ease data collection and management, Jim Comiskey and Roger Mosher developed the BioMon Suite, a database tool for capturing the information gathered with the MAB tree protocol. According to MAB, "BioMon frees users from worrying over database design or calculation errors. Instead, researchers can concentrate on the quality of the data collected and on interpretations of summarized information" (National Zoological Park, 2004a, Download BioMon). As described in the BioMon user manual (O'Grady, 1999, p. 7), data entry is performed according to a logical data collection pattern, first establishing reserves, plots, quadrats, and finally individual trees. Measurements are then assigned to each tree and associated with a census year. Researchers may then use BioMon to generate basic reports on the data, generate maps of the individual plots, or export the data to another format.

BioMon runs under 32-bit Windows platforms and stores its data with the Borland Database Engine in a collection of specific tables in the format used by the Borland Paradox

application (O'Grady, 1999, p. 2). These tables constitute BioMon 'documents' and are identified to the application as such by the file folder in which they exist.

BioMon breaks the data it collects into two categories: session parameters and the sample data themselves. A user configures session parameters before each collection run with descriptive information concerning the sampled data, for example the location and details of the reserve from which it originates (O'Grady, 1999, p. 9). Some of these fields, such as longitude and latitude of the plot, are structured, but many consist of long freeform text fields to capture anything that may be relevant. At each level, a code is assigned to the session data being entered so that it may be referred to on other entry screens or reused at later dates. Figure 1 shows the screen for capture of plot session parameters, which is representative of the other session data screens. This screen shows entry for the plot coded "2" assigned to the session data coded "DC", visible in the upper left corner.

|                        |                                       |                      |                                     |  |   |
|------------------------|---------------------------------------|----------------------|-------------------------------------|--|---|
| ResCode                | <input type="text" value="DC"/>       | minDBH (cm)          | <input type="text" value="1.000"/>  | <input type="button" value="Clear"/>   | <input type="button" value="Undo Changes"/>   |
| Plot No.               | <input type="text" value="2"/>        | Orientation          | <input type="text" value="0.000"/>  | <input type="button" value="Add New"/> | <input type="button" value="Save Changes"/>   |
| Description            | <input type="text"/>                  | Altitude             | <input type="text" value="0.00"/>   | <input type="button" value="Delete"/>  |   |
| Lat (decimal degrees)  | <input type="text" value="0.000000"/> | Plot Width (meters)  | <input type="text" value="100.00"/> | <input type="button" value="Prev"/>    | <input type="button" value="Next"/>   |
| Long (decimal degrees) | <input type="text" value="0.000000"/> | Plot Height (meters) | <input type="text" value="100.00"/> | Quad Type                              | <input type="text" value="S"/> S: standard plots<br>R: non-standard plots<br>C: circular quadrats |
|                        |                                       |                      |                                     | SiteCode                               | <input type="text"/>  |

|          |                      |
|----------|----------------------|
| Abstract | <input type="text"/> |
| Projects | <input type="text"/> |
| History  | <input type="text"/> |
| DataSets | <input type="text"/> |
| Findings | <input type="text"/> |

**Figure 1: Entry of plot session parameters in BioMon**

(Source: BioMon, 1999)

After entering session data, the user may progress to entering the measurements taken in the field (O’Grady, 1999, p. 15). This occurs in two stages, first adding all trees to be sampled with position, species, and description in the upper portion of the screen, and then assigning periodic measurements such as DBH and height to each tree in the lower portion. Figure 2 shows the relevant screen for this process; the upper half of the screen is used for adding trees and the lower half is used for attaching samples.

|  |                                |   |   |  |                                |
|--|--------------------------------|---|---|--|--------------------------------|
| <b>Session Parameters</b>  |                                | Reserve                                     | <input type="text" value="DC"/>                       | Plot No                                      | <input type="text" value="2"/> |
|  |                                | Census Year                                 | <input type="text" value="2004"/>                     | Quad No                                      | <input type="text" value="1"/> |
| <b>Tree Information</b>  |                                | Tree Position: Surveyed                     |   | Tree Position: Stored.                       |                                |
| Tree No  | <input type="text" value="1"/> | Line  | <input type="checkbox"/> A <input type="checkbox"/> B | X  | <input type="text"/>           |
| Species  | <input type="text"/>           |   |   | Y  | <input type="text"/>           |
| Date Recorded  |                                | Tree Description:                           |   |  |                                |
| <input type="text" value="09/14/2004"/><br><small>(mm/dd/yyyy)</small> |                                | <input type="text"/>                        |   |  |                                |
| <input type="button" value="Add New Tree"/>                            |                                | <input type="button" value="Save Changes"/> |   | <input type="button" value="Undo Changes"/>  |                                |
|  |                                |   |   | <input type="button" value="Clear Form"/>    |                                |
|  |                                |   |   | <input type="button" value="Enter Samples"/> |                                |
| <b>Sample Information</b>  |                                |   |   |  |                                |
| Stem No  | <input type="text"/>           | Angle of Fall                               | <input type="text"/>                                  |  |                                |
| DBH (cm)   | <input type="text"/>           | Observation Code                            | <input type="text"/>                                  | Phenology Code                               | <input type="text"/>           |
| Height(m)  | <input type="text"/>           | Status                                      | <input type="text"/>                                  |  |                                |
| Notes  |                                |   |   |  |                                |
| <input type="text"/>   |                                |   |   |  |                                |
| <input type="button" value="Add New Sample"/>                          |                                | <input type="button" value="Save Changes"/> |   | <input type="button" value="Undo Changes"/>  |                                |
|  |                                |   |   | <input type="button" value="Clear Form"/>    |                                |
|  |                                |   |   | <input type="button" value="Edit Trees"/>    |                                |

**Figure 2: Entry of sample data in BioMon**

(Source: BioMon, 1999)

After completing data entry, BioMon can produce reports in different formats, including maps which are used for field verification. BioMon formats these reports as simple text files, easily read by researchers for analysis or by computers for further processing.

### 3.2.4 SALVIAS

Based at the University of Arizona, Project SALVIAS (Synthesis and Analysis of Local Vegetation Inventories Across Scales) is an Internet-based collection and distribution point for vegetation data collectors. Originally inspired by Alan Gentry, the project now

covers most of Central and South America (University of Arizona, 2004, Mission Statement). In SALVIAS, twenty-four participants from ten organizations are able to submit, maintain, and allow other members access to their submitted data at their discretion. This database provides no means to analyze data, only access and a medium for distribution.

#### ***3.2.4.1 Mission***

The intention of SALVIAS is to allow the collaboration of researchers through the distribution of plot taxonomy and herbarium databases. Its specific goals are to:

- “1. Provide a central repository for plant community inventories
2. Compile links to databases of taxon-level and geophysical attributes
3. Develop informatics tools needed to link these heterogeneous data sources
4. Make this information available to a wide audience of potential investigators and collaborators, while simultaneously protecting the intellectual property rights of individual data providers”

(University of Arizona, 2004, Mission Statement)

All four points must be considered when establishing such a widespread database.

#### ***3.2.4.2 Membership***

Becoming a participant in SALVIAS (University of Arizona, 2004, Become a SALVIAS Member) begins with the submission of an online application. Fields such as name, login name, address, and affiliated institution are examples of information one must provide the administrators of SALVIAS. Only SALVIAS members can submit and access data.



### ***3.2.4.3 Data submission***

Minimal effort is required on the part of the user to submit data to SALVIAS. Understandably, different data collectors from around the world have different standards by which they organize their data. SALVIAS therefore has no set formatting standards for data submission; contributors are allowed to send in information as-is (University of Arizona, 2004, Information for Data Contributors). Also, SALVIAS is capable of retrieving database information from a contributor's website if the contributor provides Internet access to their database.

### ***3.2.4.4 Security and access***

SALVIAS members may access any researcher's submitted data (University of Arizona, 2004, Intellectual Property Rights Policy). Types of data include plot data and herbarium data. Plot data are sampled from specific plot areas around the world. Herbarium data are sampled from particular groupings of plants established, maintained, or otherwise monitored by an organization. Varying degrees of security are needed for each type of data. SALVIAS has established an Intellectual Property Rights Policy for its members. By virtue of being a member of SALVIAS, one is subject to accessibility and information-release terms. Table 1 shows an example of the description data available to full-access members.

**Table 1: SALVIAS full-access sample description data**

(Source: University of Arizona, 2004, Intellectual Property Rights Policy)

| <b>Field Name</b>          | <b>Example</b>                        |
|----------------------------|---------------------------------------|
| Plot Name                  | e1750-3                               |
| Plot Admin Name            | Brad Boyle (bboyle@email.arizona.edu) |
| Project Name               | Boyle Transects                       |
| Project Admin Name         | Brad Boyle (bboyle@email.arizona.edu) |
| Country                    | Ecuador                               |
| State/Province             | Carchi                                |
| Latitude                   | 0.88                                  |
| Longitude                  | -78.16                                |
| Elevation                  | 1750 m                                |
| Holdridge Life Zone        | Tropical Lower Montane Pluvial Forest |
| Mean Annual Temperature    | 17.9                                  |
| Total Annual Precipitation | 5000                                  |
| Area                       | 1000 m                                |
| Size Minimum               | 2.5 cm dbh                            |
| Method Notes               | Modified Gentry tenth ha. transect    |
| Method Reference           | Boyle, B 1996                         |
| Vouchered?                 | Yes                                   |
| Associated Data            |                                       |

The information SALVIAS (University of Arizona, 2004, Intellectual Property Rights Policy) provides users includes plot data and metadata. Certain fields of these data can be sensitive to an area's existence for various reasons. Metadata include location, date, plot type and plot administrator information. Plot data are the raw information collected by the field researchers such as tree cover, species, and tree density. In order to ensure the integrity of biologically diverse plots from interference by outside parties, several security restrictions are in place concerning the project administrator, plot administrator, and the user of the data

(See Table 2). The “user”, in general terms, can then either be completely restricted, only allowed to view the plot description, or be granted full access to all data.

**Table 2: SALVIAS access permissions**

(Source: University of Arizona, 2004, Intellectual Property Rights)

| Level          | Grants Permission to... |            |               |                |
|----------------|-------------------------|------------|---------------|----------------|
|                | User                    | Plot Admin | Project Admin | Database Admin |
| User           | False                   | False      | False         | False          |
| Plot Admin     | True                    | False      | False         | False          |
| Project Admin  | True                    | True       | False         | False          |
| Database Admin | False                   | True       | True          | True           |

Unlike plot data, herbarium data are not subject to any access or security restrictions; they are freely accessible among SALVIAS members (University of Arizona, 2004, SALVIAS Data Sources). Herbarium data come from a secure location and therefore, it is considered that no added layer of protection is needed. Provided one is a participant of SALVIAS, one has unrestricted access to herbarium data.

### ***3.3 Data Processing***

Through analysis, data become useful in informing decision-makers in biodiversity management. Once cleaned, statistics taken across plots and regions can be utilized to gain an understanding of ecological and other factors which affect biodiversity both positively and negatively. The information gained from this study will better equip decision-makers to evaluate policies which affect biodiversity.

### 3.3.1 Cleaning Data

Meaningful comparisons may only be made between data from different plots if taxonomic names are consistently applied. Variation in spelling or nomenclature may lead to counting trees of the same species as distinct species and thus lead to inaccuracies in statistics. Therefore, cleaning data sets to ensure consistency in naming is an essential first step to the analysis process. The TROPICOS database and TaxonScrubber are two resources available to ease this task.

#### 3.3.1.1 Taxonomic Naming Conventions

The taxonomic hierarchy is organized into levels starting at kingdom and descending through division, class, order, family, genus, and finally species. The last three components, family, genus, and species, are those recorded in the field and the most pertinent in examining plot data. For example, the species commonly referred to as white oak is of the family Fagaceae and the genus *Quercus*, and the species *alba*.

Often trees cannot be identified to a precise species in the field but can be narrowed to a genus. In this case, they may instead be identified by a morphospecies, a unique identifier provided by the field researcher in place of a specific epithet. All trees which the collector can identify as belonging to the same species within that plot will be catalogued under that identifier, though their specific species may be unknown. These identifiers must only be considered valid for comparisons within one plot, however; no guarantee is made that species A of some genus in one plot is the same as species A of the same genus in any other plot.

In other cases, the researcher may simply mark the species as unknown, making no guarantees that all trees marked as unknown belong to the same species. Occasionally only

the family may be identified and the genus must be marked as unknown. More rarely, a tree cannot even be identified to the family level and must be recorded as completely unknown.

### **3.3.1.2 TROPICOS**

The Missouri Botanical Garden, located in the city of St. Louis, is a center for botanical research and science education (Missouri Botanical Garden, 2004, About MBG). It contains bryophyte, vascular cryptogam, gymnosperm, monocotyledon, and dicotyledon specimens. Aside from its private collection, the Garden provides a taxonomic tool called W<sup>3</sup>TROPICOS through its website. Jim Solomon, curator of the Garden, maintains this site. TROPICOS enables anyone to search and reference a complete list of plants by scientific name. Group, family, genus, and species are given along with published article references for name verification (Missouri Botanical Garden, 2004, TROPICOS). TROPICOS allows users to search existing scientific names and will not return results for incorrect spellings or nonexistent queries. Therefore the TROPICOS database ensures consistent naming of species between groups of researchers. Herbarium specimen pictures are also provided for visual identification of a specific inquiry.

### **3.3.1.3 TaxonScrubber**

SALVIAS (University of Arizona, 2004, SALVIAS TaxonScrubber) also produces a cleaning tool called *TaxonScrubber*. *TaxonScrubber* is able to sort through a provided list of species names, compare them to the standard taxonomic names as listed in the TROPICOS database, and provide suggested corrections. Recommended corrections to names may not always be provided, but the field will be flagged for manual review. A user may then navigate a species list and interactively clean the data (See Figure 3).

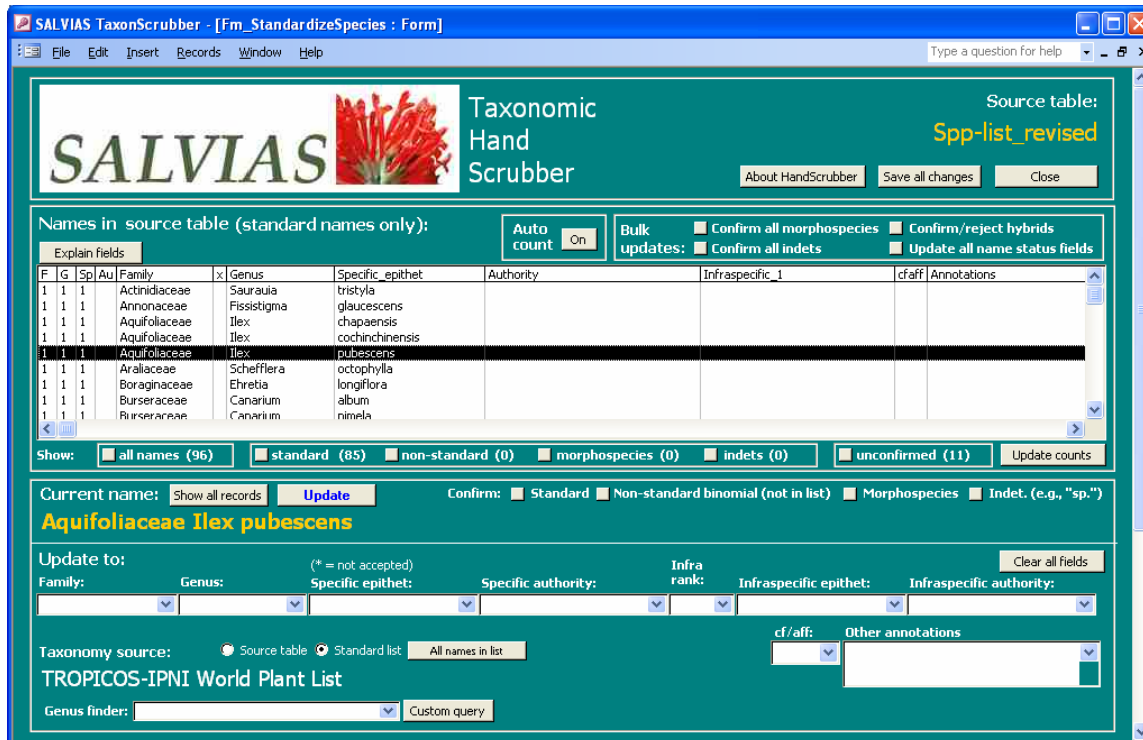


Figure 3: TaxonScrubber interface

(Source: TaxonScrubber)

### 3.3.2 Statistical Analyses

In order to attain a more general view of biodiversity within a sampled region, MAB hopes to compute numerical measures such as stem density and basal area to compare distribution of species and families between plots (A. Alonso, personal communication, October 26, 2004). Basal area of a species is defined as the sum of the cross-sectional areas at breast height of all trees of that species, expressed in square meters per hectare. Stem density is the number of trees measured in the region, expressed as stems per hectare. In addition to these measures, the number of quadrats in which individuals of a species appear, called the species frequency, serves as a heuristic representation of how clustered or widespread a species appears. Relative basal area, relative stem density, and relative

frequency may then be computed for a species as the fraction of the total basal area, total stems, and total frequency in a plot which that species accounts for. The three relative measures are without units.

By computing these measures on species in a plot, one can form hypotheses about the state of the plots. For example, a high stem count with lower basal area might indicate new growth after recent cutting while a low stem count with a high basal area might indicate much older growth. By adding these three figures (relative stem density, basal area, and frequency), one can generate importance value indices (IVI) to rank how important species are within the plot. The entire process may be repeated upon families, rather than species, to determine how important families are within the plot.

These measures, therefore, aid researchers in obtaining a greater understanding of the factors at work within a region by providing a numerical basis through which plots may be compared. BioMon is equipped with reporting tools which can calculate these statistics for each plot census. It generates reports such as that seen in Table 3 that may be regarded as a summary of the plot.

**Table 3: First 10 rows of BioMon statistics for Luquillo National Forest**

(Source: BioMon)

| SpCode | Total<br>Trees | Total<br>Stems | Average<br>DBH (cm) | Basal<br>Area | Relative<br>Density | Relative<br>Dominance | Relative<br>Frequency | Relative<br>IVI |
|--------|----------------|----------------|---------------------|---------------|---------------------|-----------------------|-----------------------|-----------------|
| PREMON | 136            | 136            | 15.646              | 2.615         | 30.91               | 9.33                  | 13.37                 | 53.61           |
| DACEXC | 40             | 40             | 47.526              | 7.096         | 9.09                | 25.32                 | 9.63                  | 44.04           |
| SLOBER | 83             | 83             | 21.952              | 3.141         | 18.86               | 11.21                 | 11.76                 | 41.84           |
| INGFAG | 38             | 38             | 25.773              | 1.983         | 8.64                | 7.07                  | 8.56                  | 24.27           |
| GUAGUI | 20             | 20             | 46.164              | 3.348         | 4.55                | 11.95                 | 6.42                  | 22.91           |
| SAPLAU | 17             | 17             | 32.652              | 1.423         | 3.86                | 5.08                  | 5.88                  | 14.83           |
| DIDMOR | 17             | 17             | 31.627              | 1.336         | 3.86                | 4.77                  | 5.35                  | 13.98           |
| ALCFLO | 11             | 11             | 47.789              | 1.973         | 2.50                | 7.04                  | 4.28                  | 13.82           |
| MANBID | 12             | 12             | 38.788              | 1.418         | 2.73                | 5.06                  | 4.81                  | 12.60           |
| CECSCH | 8              | 8              | 40.064              | 1.009         | 1.82                | 3.60                  | 3.21                  | 8.63            |

### 3.3.3 Java for Manipulating Data

Software applications such as Microsoft Excel and Access are sufficient for straightforward tasks involving data, such as sorting, manipulating columns, or performing basic mathematical operations. Some tasks, though, are more complicated, particularly time-consuming, prone to human error, or repeated frequently. For these tasks, small computer programs created with Sun Microsystems' (Sun, 2004) Java programming language are often more capable tools. Programs tailored to the task allow processes to be automated and thus performed more quickly with less work. The Java language enables rapid development of these tools and provides a stable platform, as "any computer system with the Java VM installed can run a Java program regardless of the computer system on which the application was originally developed"(Java Technology Overview). Java is a high-level language and



provides a large set of tools through the Java Application Programming Interface (API) for manipulating data and graphically interacting with users.

## **4 Methodology**

This project aimed to improve the gathering, storage, and distribution of sampled tree data in the Monitoring and Assessment of Biodiversity Program. To accomplish this, we first focused on data from North America, a selection made under the guidance of our liaisons after an initial survey of the data situation, to generate a process which could be generalized for use on the other regions. The team performed the steps needed to organize the sample data for that region and documented the process so that others could repeat it for any data submitted to the MAB in the future. We then applied it to the remainder of the plot data held by MAB. We approached managing MAB's data from three different perspectives. First, we organized and imported the data that MAB had already collected into BioMon. We then identified scientists for potential future plot data sources to increase the volume of data in MAB's possession. Finally, we prepared summaries of the data sets for upload to the web and packaged the newly organized data sets themselves for transmission to the SALVIAS network.

### ***4.1 Improvement of Internal Access to Plot Data at MAB***

Improving data management at MAB demanded that we first determine their current status. To this end, the team conducted an inventory of the data held on computers and media in the MAB offices. This included examination of the files that contain the data, including BioMon databases, and review of all documentation that MAB kept on them, including the internal summary of their holdings. Through this process, we generated an index that accounted for present sets of data and identified all missing sets. Alongside this index of existing data, the team compiled a metadata inventory for each data set, including where and

when it was collected, who was responsible for its collection, and the format and location of the data themselves. The resulting table enabled the project team and MAB to more easily determine the status of data in storage and provided guidance for taking future steps regarding plot data management.

#### **4.1.1 Organize Data into Regions**

The first step the project team took towards accomplishing the goal of improved internal access was the organization of the plot data present at MAB. A summary list with all the known sources of potential plot data had already been produced by MAB. It listed the site, country, number of plots, data availability, year initiated, researcher, latitude, longitude, minimum diameter at breast height (DBH), number of trees and species within the plot, and the file directory. Upon further examination, this list was found to be out of date with respect to the data in MAB's possession. A bottom-up approach to creating an inventory of MAB data had to be taken instead.

The resulting inventory of MAB's data helped the team determine further steps in the data organization process. The inventory contained information about the region, country, number of plots, year of census or re-census, file format and location, and a checklist for the components of the data. Table 4 shows the North America portion of the table; the complete document is included as Appendix B.

**Table 4: Except from inventory, North American plots**

| <i>Region</i> | <i>Country</i> | <i>Plot #</i> | <i>Year</i> | <i>Re-census?</i> | <i>Format</i> | <i>File directory</i> | <i>Files exist [y/n?]</i> | <i>Species list [y/n?]</i> | <i>Sample data [y/n?]</i> |
|---------------|----------------|---------------|-------------|-------------------|---------------|-----------------------|---------------------------|----------------------------|---------------------------|
| NA            | Canada         | 1             | 1994        |                   | csv           | Backup\Kejimkujik     | Y                         | Y                          | Y                         |
| NA            | Canada         | 2             | 1994        |                   | csv           | Backup\Kejimkujik     | Y                         | Y                          | Y                         |
| NA            | Canada         | 1             | 1995        |                   | csv           | Backup\Long Point     | Y                         | Y                          | Y                         |
| NA            | Canada         | 1             | 1995        |                   | csv           | Backup\Rocky Point    | Y                         | Y                          | Y                         |
| NA            | Canada         | 1             | 1999        |                   | biomon        | Royal Roads Un        | Y                         | Y                          | Y                         |
| NA            | Canada         | 2             | 1999        |                   | biomon        | Royal Roads Un        | Y                         | Y                          | Y                         |
| NA            | U.S.           | 1             | 1993        | Y                 | csv           | Backup\Front Royal    | Y                         | Y                          | Y                         |
| NA            | U.S.           | 1             | 1995        |                   | csv           | Backup\Front Royal    | Y                         | Y                          | Y                         |
| NA            | U.S.           | 2             | 1993        |                   | csv           | Backup\Front Royal    | Y                         | Y                          | Y                         |
| NA            | U.S.           | 3             | 1994        |                   | csv           | Backup\Front Royal    | Y                         | Y                          | Y                         |
| NA            | U.S.           | 4             | 1994        |                   | csv           | Backup\Front Royal    | Y                         | Y                          | Y                         |
| NA            | U.S.           | 5             | 1995        |                   | csv           | Backup\Front Royal    | Y                         | Y                          | Y                         |
| NA            | U.S.           | 6             | 1995        |                   | csv           | Backup\Front Royal    | Y                         | Y                          | Y                         |
| NA            | U.S.           | 8             | 1996        |                   | csv           | Backup\Front Royal    | Y                         | Y                          | Y                         |
| NA            | U.S.           | 9             | 1997        |                   | csv           | Backup\Front Royal    | Y                         | Y                          | Y                         |

Data were found which the previous summary sheet had marked “unreceived”, some data were not included on the previous summary sheet, and others were listed on the summary sheet but could not immediately be located. The summary document, including details regarding census years and plot numbers, was brought up-to-date. Relative to the amount of data MAB possessed, there were comparatively few disagreements between the summary list and the new inventory of data present.

Once this step was complete, a simple data sort was performed to organize the data by geographic region. The totals for each region can be seen in Table 5. The regions were assigned codes and added in a column to the inventory as North America (NA), Africa (AF), Asia (AS), South America (SA), and the Caribbean (CR). MAB then recommended that South America and the Caribbean be grouped together for organizational purposes. From there the data were ready to be imported to BioMon, grouped by geographic region (continent).

**Table 5: Distribution of MAB censuses by continent**

|               |     |
|---------------|-----|
| North America | 15  |
| South America | 82  |
| Africa        | 28  |
| Asia          | 4   |
| Total         | 129 |

#### **4.1.2 Test Run of Data Movement**

Before any further data processing could be performed, a test run of the planned organizational process was implemented to ensure the procedure would generate results as planned. Due to the large size of MAB’s data collection, working with a limited subset allowed us to more easily correct and repeat a procedural step if necessary. The group of data under North America, as seen Table 6, was selected for this test due to its relatively manageable size and variation of format: fifteen sets of data in both CSV (comma separated values) and BioMon formats. Once the data for North America were consolidated into a single BioMon database with proper taxonomic names, the data were put aside until the data from South America (ninety-nine sets of data, both the bulk of MAB’s collection and their main area of interest) were also processed up to this point. North and South America were then carried through the remainder of the organization step to complete exhaustive documentation of the process. That documentation was then tested by applying its instructions to the Africa and Asia data.

**Table 6: Distribution of MAB censuses by region and format**

|               | BioMon    | CSV       | Access    | Paradox   | Total      |
|---------------|-----------|-----------|-----------|-----------|------------|
| North America | 2         | 13        | 0         | 0         | <b>15</b>  |
| South America | 37        | 5         | 13        | 27        | <b>82</b>  |
| Asia          | 0         | 4         | 0         | 0         | <b>4</b>   |
| Africa        | 20        | 8         | 0         | 0         | <b>28</b>  |
| Total         | <b>59</b> | <b>30</b> | <b>13</b> | <b>27</b> | <b>129</b> |

### 4.1.3 Move Data Formats to BioMon

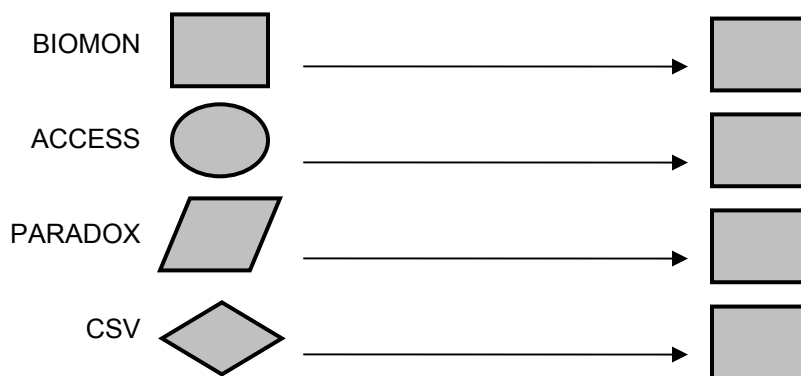
Plot data MAB was known to have in its possession existed in many different forms. These included Microsoft Access tables, Excel spreadsheets, comma separated values, BioMon export files (.bak), and Paradox database tables. The distribution of formats is shown in Table 7.

**Table 7: Distribution of MAB censuses by data format**

|         |     |
|---------|-----|
| BioMon  | 59  |
| CSV     | 30  |
| Access  | 13  |
| Paradox | 27  |
| Total   | 129 |

If a data set already existed as a BioMon file, it was simply verified as being complete and moved to the appropriate folder location by region for the next step. It required some work to move data in other formats to a BioMon-compatible CSV form for import, generally

involving adjustment to the columns in the data and, for the more complex Access and Paradox formats, the conceptual layout of the data (See Figure 4). For future reference, we recorded in detail how different types of files reach this state (See Appendix C, “Steps for Data Conversion”).



**Figure 4: First step in merging, converting all data to BioMon format**

#### **4.1.4 Harmonize Taxonomic Names, Authorships, and Codes**

When collecting tree data in the field, researchers use BioMon as a convenient method for organization and storage of their collected data. Part of the data entry process requires the researcher to identify the species of a tree. In some instances, taxonomic names for the appropriate tree were misspelled or based on a different standard, a minor issue in a single database where the variation would be consistently enforced but a major complication when conducting comparisons between databases. Also, BioMon allows researchers to pick codes for their species which are convenient to their work. Ultimately, though, this flexibility leads to a variety of code sets which further complicates comparisons between surveys. We therefore unified the species codes and the family, genus, and species of the trees under the TROPICOS standard naming convention and consistently enforced those codes upon all the

data now present in BioMon. Finally, the authorship of a species (the first person to document and describe a species) was recorded within BioMon to provide a source for the species name.

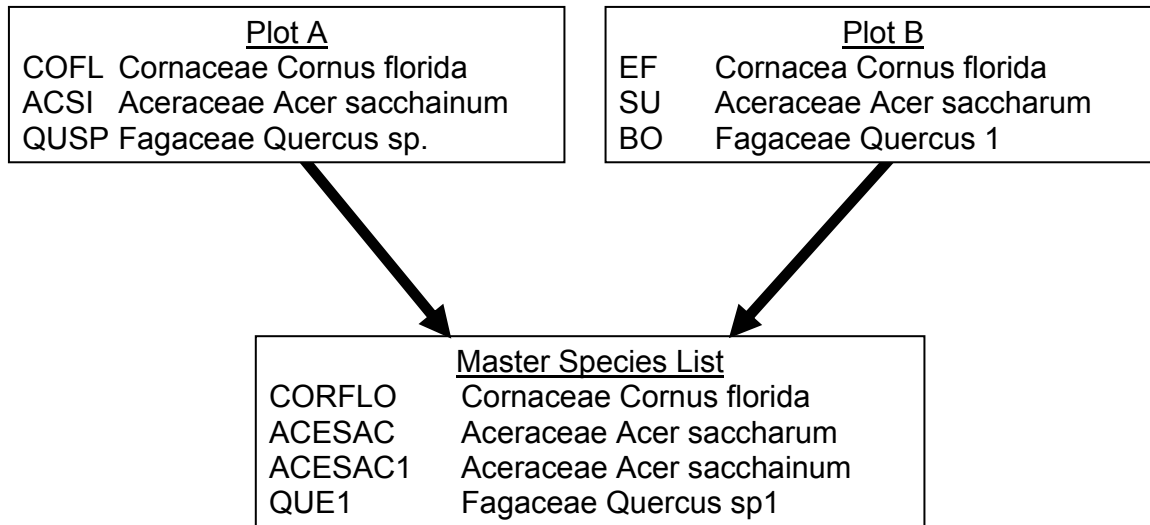
#### ***4.1.4.1 Names and Authorships***

Starting at the site level, BioMon was used to export the species list as comma separated values. Then the list of species for each site was reviewed to ensure a valid taxonomic name and authorship for each species. If one of these was in error or missing, it was entered correctly into the file according to information that was found within the TROPICOS database.

#### ***4.1.4.2 Species Master List with Codes***

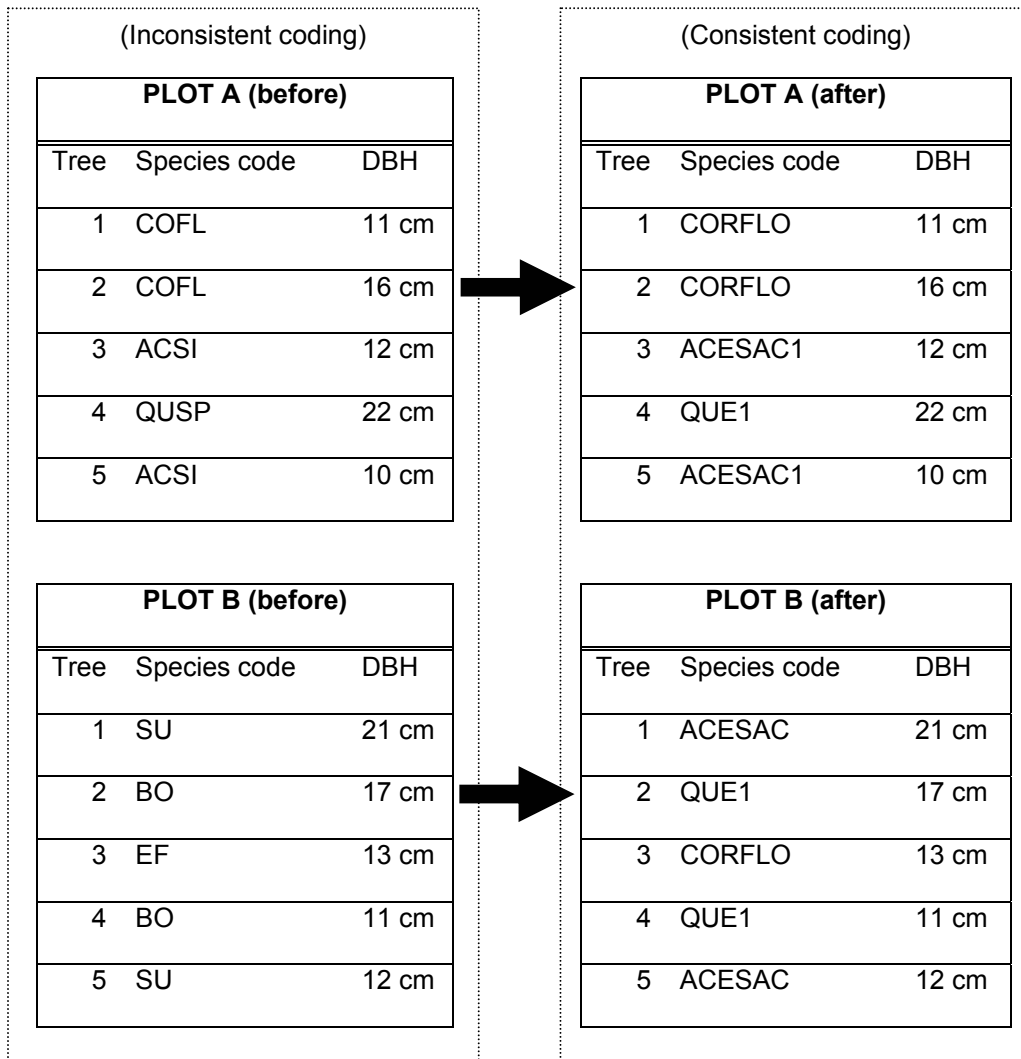
Multiple sites' species lists were then combined into a master species list for the entire region as demonstrated in Figure 5. Note that the spelling error in "Cornus florida" in Plot B has been corrected and, though it appears in both plots, it has only one entry in the master list. Also note the naming convention for the Quercus morphospecies has been standardized, though an individual reading these data must be aware that trees in Plot A labeled as 'QUE1' are not necessarily the same species as those in Plot B labeled 'QUE1'.





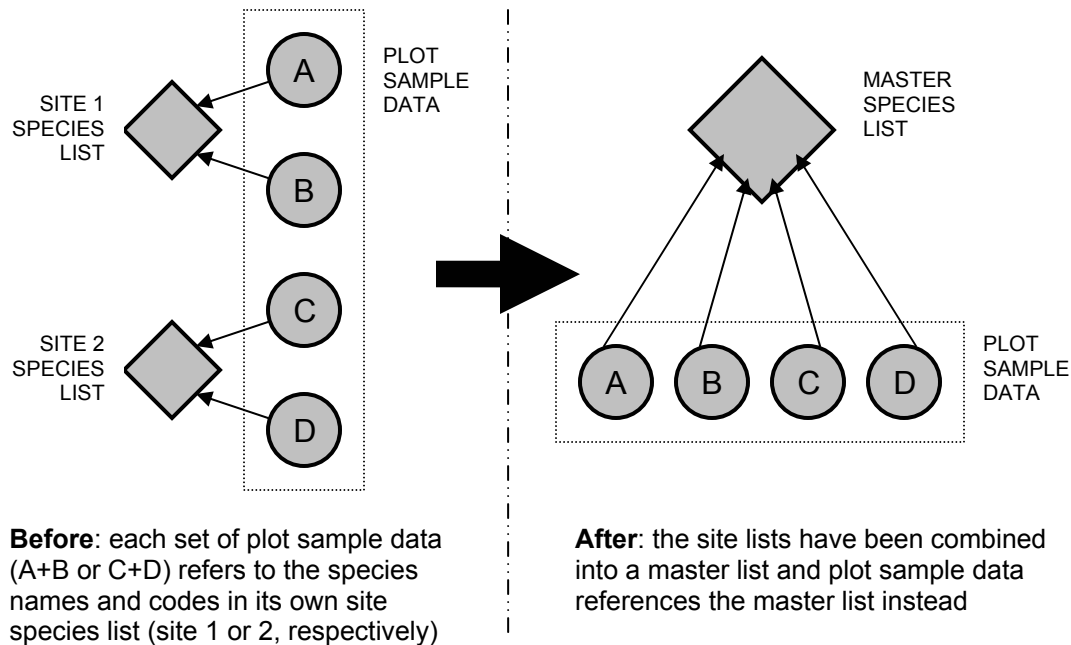
**Figure 5: Simplified example of combining species lists**

Once this process is complete, the old sample data is ready for adjustment to reflect the new list by replacing the old codes from their old species lists with the new master codes from the new consistent species list (see Figure 6 for an example). The new codes were generated from the first three letters of the genus and the first three letters of the species with additional digits to enforce uniqueness when required. Morphospecies were coded as the first three letters of the genus and the number used by the researcher when entering the data. In the case where two genera have the same three initial letters, an additional character was appended after the number.



**Figure 6: Simplified example of adjusting species codes for plots in previous figure**

Importing the resulting updated samples into a single regional BioMon database brings the process full circle, as shown in Figure 7.



**Figure 7: Before and after species list organization**

#### **4.1.4.3 Automation**

Because this task necessitated examining several thousand species, a Java program was produced to automate the task of looking up the species, verifying it, identifying its authorship, and assigning it a new species code. Human involvement was therefore only necessary in the instances where the species name was not found listed in TROPICOS, generally because of a typographical error in data entry. Using a program greatly reduced the amount of time required for the task, as approximately four out of every five species entries were already correct and therefore did not need to be examined by hand. In the future, this process will have to be repeated for incoming plot data and as such was documented by the project team (see Appendix C).

#### **4.1.5 Produce Summary Data**

The project team then moved on to produce summary analyses (see Appendix D) and graphical representations for each plot. These summary data contain lists of the ten most prevalent species and families in a plot by importance value index (IVI) and family importance value (FIV) respectively. Along with the summary information, a graph for each plot was constructed for better visual comparison of species IVI and population in an area. Originally Access queries generated statistics from the CSV forms of the data which were used, along with graphing tools, to summarize information in a word processor. However, as this proved prohibitively labor-intensive given the size of the plot collection, another Java program was produced to automate the task. This program used the raw CSV data and metadata tables extracted from Biomon to generate the final reports directly in PDF form, including all cover sheets, page headers, data tables, and graphs.

#### ***4.2 Increase the Quantity of Plot Data MAB Holds***

In order to receive additional information recorded from vegetation plots, the team established contact with the scientists who have access to that data. Referring to MAB's data summary list and the inventory list created by the project team, two instances of potential data sources existed: sites recorded as established but from which data had not been sent to MAB and sites for which MAB has some data but for which additional data may exist. Additionally, MAB taught several courses on how to develop biodiversity plots and conduct a census in the field. Potentially, some of the participants of those courses had followed through and established plots. We contacted a majority of the participants of those courses as well. The appropriate contacts have thereby been established to obtain vegetation plot data from the sources identified through MAB resources.

### **4.2.1 Sites with Previously Submitted Data**

According to MAB's previous summary list and the team's inventory list, there were one hundred sixty-six plots from thirty-nine sites around the world which have previously sent data to MAB. Through web-based research, review of papers associated with the plots, and interviews with MAB staff who have worked with the plots, the team identified the primary points of contact for each plot and established contact information for those individuals. Originally, we had intended to contact these individuals with a mailing prepared alongside MAB Assistant Director A. Alonso (see Appendix E). Upon further consideration, however, MAB opted to contact these individuals informally through the channels of normal business. The contact information was therefore recorded in BioMon and reported to MAB.

### **4.2.2 Planned Sites without Submitted Data**

The team also identified the sites that in the past signaled their intent to establish biodiversity plots to MAB yet had never submitted data. As with the previous group, a mailing was prepared (see Appendix F) and the team identified and found contact information for the individual responsible for the plot. However, rather than send out the mailing as planned, this information was likewise reported for MAB to follow-up upon through its own contacts.

### **4.2.3 SI Biodiversity Training Course Participants**

For MAB-conducted courses on establishing biodiversity plots, we took contact information from course records for each of the participants. The team then contacted these individuals through e-mail to inquire whether they had conducted data collection since their participation in the course (see Appendix G). Those who responded positively were recorded for MAB to follow-up upon with the previous two groups. For these letters, J. Sevin, project

liaison and MAB Education and Training Coordinator, granted the use of her name and e-mail address to provide enhanced credibility. At her request, the e-mail included a brief survey for evaluation of the courses these individuals participated in, but evaluation of these survey results fell outside the scope of our project.

### ***4.3 Make the Data Available to Researchers Outside of MAB***

Once the data were organized and easier to handle, the team could focus on improving the access of those outside of MAB to the plot data. This included both sending them to the SALVIAS network for distribution via its online database and publishing summary information on the MAB web site. The former made the data available to SALVIAS members in its raw form, while the latter provided a general catalogue of the data to the public at large.

#### **4.3.1 Send Data Sets to SALVIAS**

The SALVIAS webpage (University of Arizona, 2004, Information for SALVIAS Data Contributors) made apparent that organization's flexible requirements for data submission and willingness to accommodate those interested in sharing data. This proved to be the case; at the request of MAB, database administrator Brad Boyle from SALVIAS began producing a mechanism to upload data directly from BioMon databases into the SALVIAS network.

With the data already in consistent, regional databases, little additional work was required to transmit them. The respective BioMon files were compressed and prepared for transmission. At the time of this report the import mechanism at SALVIAS was still in its development phase. Through electronic transfer, MAB will easily be able to send the data to SALVIAS once their systems are ready for it.

### **4.3.2 Publish Summary Data**

Finally, to improve the visibility of the data to the general public, the summary documents generated through the organization phase were prepared for upload to the MAB web site <<http://www.si.edu/simab>>. The uploaded documents included the table demonstrating the components of the IVI and FIV calculations along with graphs for easier visual comparison. The statistics and graphs were accompanied by a brief explanation of how the reports may be interpreted and what the different calculated values represent.

### ***4.4 Materials Produced for Future Work***

Beyond the organized data and the contacts established with SALVIAS and researchers, we produced several artifacts for future use by MAB in the interests of streamlining data handling in the future. These products include software programs that clean and subsequently merge species lists and generate summary reports. Also, the process documentation we produced will serve as a manual for data handling at MAB. With these two initiatives, future work with plot data will be easier.

For the areas of the data organization procedure which were labor-intensive but very mechanical, specifically species list cleaning, species list merging, and summary production, we developed software tools in Java to ease the burden on the data handler. We also created a record of the work we performed in a data handling manual (see Appendix C). We aimed to aid any future plot data handlers at MAB in the movement of data into their local database and on to SALVIAS and the web. The manual provides step-by-step instructions on how to identify plot data sets, convert them to BioMon format, manipulate them within BioMon, unify species names between them, combine them into regions, produce and calculate summary reports for them, and submit the organized and cleaned results to SALVIAS.

Several individuals within the MAB office, BioMon users and novices among them, were selected to proofread this document to verify its accuracy and readability. Each member of our team also went through the manual with a set of data from start to finish verifying that steps had not been omitted. The team left a printout of this manual in the MAB office as well as copies in Adobe Acrobat form (PDF) on the MAB internal file server and web site.



## 5 Results and Discussion

### 5.1 Improvements to Access at MAB

As a result of this project, MAB has seen improvements to its internal ability to access its vegetation data. These include a more thorough record of which data are held, four easily accessed region databases containing all plot data, a procedure for maintaining those four databases as new data are received, and software tools to supplement that procedure.

#### 5.1.1 Complete Inventory

MAB now has a complete list of each individual set of census data it possesses, totaling 129 censuses conducted upon eighty-one plots within twenty-six reserves (see Table 8). One can infer that a total of forty-eight subsequent censuses (number of total censuses minus number of total plots) had been taken after initial plot data were collected. In addition, it has a more comprehensive list of plots which are known to exist but not represented in the data set now, including contact information and the additions gained from e-mail correspondence.

**Table 8: Distribution of MAB censuses by regional database**

| <b>Region</b> | <b>Reserves</b> | <b>Plots</b> | <b>Censuses</b> |
|---------------|-----------------|--------------|-----------------|
| North America | 5               | 14           | 15              |
| South America | 10              | 41           | 82              |
| Africa        | 9               | 23           | 28              |
| Asia          | 2               | 3            | 4               |
| Total         | 26              | 81           | 129             |

### 5.1.2 Organized Regional Databases

The data held by MAB now exist in just four regional BioMon databases, whereas before they were found in many small databases in BioMon and various other formats.

Within each regional database, the plots are separated by reserve code, properly labeled with numbers and census years, and marked with all available metadata information. As seen in Table 8, it is now much easier to locate a set of census data within a regional database.

As a component of this process, we have generated a master species list for each region which is comprehensive for all plots and their corresponding data. These lists include codes for trees identified to a morphospecies, to a genus with an unknown species, and to a family with an unknown genus, as well as complete identifications to specific epithet. Table 9 shows the frequencies of these occurrences in the master species lists. These species lists, containing complete species identifiers verified in TROPICOS, will serve as a basis for incorporating future additions to the database.

**Table 9: Code counts for each regional database for different levels of identification**

| <b>Region</b> | <b>Morpho-species</b> | <b>Unknown Species</b> | <b>Unknown Genus</b> | <b>Known Species</b> | <b>Total Species</b> |
|---------------|-----------------------|------------------------|----------------------|----------------------|----------------------|
| North America | 6                     | 0                      | 1                    | 86                   | 93                   |
| South America | 287                   | 12                     | 16                   | 988                  | 1303                 |
| Africa        | 200                   | 10                     | 10                   | 515                  | 735                  |
| Asia          | 4                     | 0                      | 0                    | 217                  | 221                  |
| Total         | 497                   | 22                     | 27                   | 1806                 | 2352                 |

### **5.1.3 Procedure for Maintaining Organization**

As we experimented through trial and error and gradually discovered how best to organize the plot data sets, we kept a precise record of how we moved from stage to stage at a level of detail another individual could use to precisely replicate our actions. This document, found in Appendix C, has been provided to MAB as a manual for the data handling process involving BioMon and describes the procedure from initial reception of data in a variety of formats, through the process of converting and merging to the regional BioMon databases, to finally preparing the summaries for the web and the polished data for SALVIAS. The manual presents the process sequentially from start to finish and goes down to the level of individual mouse actions and keystrokes. At the end of the manual, we appended solutions to technical problems which arose through our work in the hope that future lost time due to those issues may be avoided.

### **5.1.4 Software Tools to Aid Organization**

Through the organization phase of the project, we developed several software tools originally for our own use to facilitate completion of the most time-consuming portions of the organization process. These tools made possible some aspects of the project that otherwise would have proven impossible due to time constraints, such as completing authorship information and generating master species lists. Recognizing their potential usefulness to others carrying out this process in the future, we produced graphical user interfaces for these tools and documented their use for the benefit of individuals adding data in the future.

First, we developed the “BioMon Species List Cleaner” to verify species names in a BioMon species list by comparing them against the TROPICOS database and to



from various sources in a word processor, the time required to generate summaries is reduced from hours to only minutes.

These tools along with their source code are attached as Appendix H.

## ***5.2 Additional Data for MAB***

We have provided a short-term boost to MAB's data gathering efforts by identifying individuals who seemed likely to possess plot data and requesting access to their data. These groups included participants from past biodiversity training courses offered by MAB and researchers who have submitted data in the past or otherwise made it known they intended to collect data.

### **5.2.1 Course Participant Contacts**

We attempted to contact via e-mail 161 individuals who between 1997 and 2004 took , a MAB-sponsored biodiversity training course which detailed the protocols for establishing one-hectare biodiversity plots. Fifty-four of the messages were immediately returned to sender, with an eventual total of 74 as undeliverable, a problem expected because of the age of some of the contact information. Twenty-two people ultimately responded; of that group, six reported they had established or otherwise been associated with vegetation plots. We recorded these individuals alongside the researcher contacts which we established for future work in collecting plot data.

### **5.2.2 Researcher Contacts**

Contacting researchers connected to MAB who in the past had submitted data, or signaled their intent to do so, proved difficult due to the different ways in which each was connected to the MAB program. Because of the highly diverse nature of MAB's relations to

these individuals and the different missions of the reserves they are associated with, MAB requested that rather than directly contact these people, we report on the researchers who are doing plot work along with their contact information so that MAB might request their data through already established channels.

### ***5.3 Increased External Availability***

There are now the foundations for two levels of access to MAB's data sets from the outside world. The general public looking for simple vegetation data analysis may view the summary data on the MAB web site. Researchers interested in performing their own analyses and who belong to organizations which are SALVIAS members will be able to obtain immediate electronic access to the raw data through the SALVIAS web site at <<http://www.salvias.net>>.

#### **5.3.1 Web Access to Summaries**

Available for general viewing, the summary data compiled for all plots in each of the reserves were submitted in PDF form for publishing on the MAB web site at <<http://www.si.edu/simab>>. These allow individuals who are not necessarily able to analyze the raw data but who still have interest in the data (e.g., students, teachers, conservation groups, and interested industries) to gain an overview of the situation in each plot. A brief description of the analyses performed on the summaries and how to interpret them is included alongside the summary data themselves. If a reader desires the raw data, they may request it from MAB directly.

### **5.3.2 SALVIAS Access to Data**

The plot data were compressed and prepared for electronic transmission in BioMon format to be translated to the format used internally by SALVIAS, pending the completion of the import procedure by SALVIAS developer Brad Boyle. These data sets will soon be available via the web in their complete form to all organizations which are members of the SALVIAS network.

## **6 Conclusions**

### ***6.1 Regarding Internal Data Management***

Combining plot data sets from a single region together in a single BioMon database facilitates faster access to the data and provides a central record of the plot data MAB possesses. If data are adjusted for internal consistency in taxonomy and codes when combined, comparisons may be conducted between the data sets by calculating summary statistics at species and family levels. These summary statistics also aid in identifying plots by their composition in the interests of further study.

Following a defined process, like that outlined in the data handling manual, ensures consistency in data storage and lessens the time required for organization. A standard procedure for tracking metadata alongside sample data further magnifies the gains observed from centralizing data by region by attaching contact information for researchers directly to the data for which they are responsible.

Much of the process of data organization stands to benefit from software assistance. MAB has begun this process by adopting BioMon for data entry. Additional aspects of the process which are well-defined, often repeated, and labor intensive also stand to gain from tailored software tools. This includes the processes of cleaning species lists, merging sets of data from different sites, and generating summary information.

The process of organizing data at MAB does not end with this project and must continue as MAB continues to receive new data. All systems of organization, processes, software tools, and decisions regarding data management must therefore address the continually growing nature of the MAB data set.



## ***6.2 Regarding Increasing the MAB Plot Collection***

Much data exist already collected and in an electronic form which MAB could obtain to increase its collection. In addition, there are ongoing resurveys of past plots and many new data collection plots planned for the near future. Those contacted who reported involvement in establishing and surveying biodiversity plots expressed interest in contributing to the efforts of MAB. MAB will likely find many willing contributors if it approaches people such as participants of its courses who demonstrated an interest in biodiversity and thus have potentially collected plot data. MAB will also likely find that contacting individuals and organizations which have contributed in the past will yield information about projects and plots conducted since they last contributed data.

## ***6.3 Regarding External Data Availability***

Previous data contributors and course participants were contacted and informed of recent efforts to release plot information to the SALVIAS research network. Through the SALVIAS network, MAB can provide raw data to other researchers who have use for the extreme detail included in a full set of plot data. The general public also expressed an interest in seeing the data and in participating. This transfer may be accomplished through the BioMon format itself if the data are kept organized and tagged with appropriate metadata.

Summaries on the web will reach members of the general public who are interested in the general patterns in vegetation distributions as measured by MAB plot data. This includes teachers who would like to use real data in their classes to generate interest from students, students conducting their own research, and other ecological researchers who are not interested in the level of detail in the full data set. The summaries may be generated in a

largely automated fashion from the data present in BioMon and uploaded to the web in an easily readable file format.

## 7 Recommendations

Based upon the conclusions and three goals set by the team regarding internal access to plot data, access to additional plot data, and the distribution of such data, we have produced a set of recommendations to further pursue these goals. Each builds on the work completed by the team with the intentions of magnifying gains achieved during our project.

### *Incorporate new data sets as they are received*

Specifically:

- **Track metadata with plot data**
- **Implement processes described in new data handling manual**
- **Investigate software improvements that may further facilitate the process**

Full utilization of MAB's plot data entails the proper organization and method for handling such data. As new sets of data are received, the proper steps should be taken to incorporate the data in the existing database. These steps include tracking metadata with the plot data, implementing the process described in the data handling manual, and investigating software improvements that may further facilitate the process.

By tracking metadata in BioMon with its respective plot data, valuable information such as its location, minimum DBH recorded, and researcher contact information may be identified more easily when reviewing the plots. This would require that MAB collect the information when plot data are submitted or notify researchers in the field to incorporate the information into their data entry process. The following items of information should be recorded:

- Name of researcher
- Email address
- Organization
- Department within the organization
- Telephone number
- Plot (reserve) name
- Country
- Plot location (latitude, longitude)
- Plot number
- Initial or additional plot census
- Year of census
- Size of plot (ex: 1 ha.)
- Minimum DBH of plot

With this information in hand, data sets can be fully identified, properly credited, and easily tracked by MAB and other potential data users.

Once the proper identifiers have been attached to plot data, the data should go through the process described in the data handling manual to ensure the data are identified, converted, cleaned, imported into BioMon, summarized, and sent to SALVIAS through the step-by-step directions. There may be alternative methods to those described in the manual; however, we recommend they only be explored after one becomes comfortable with the data handling process as already defined.

Having worked with the BioMon program extensively, the team has determined possible changes or updates of value to the user. These include an improved user interface, using a more modern database management system, better facilities for queries upon the data (such as with SQL), built-in facilities for moving data between databases without the extra steps currently required, and more powerful import/export features including the import of metadata. All of the features listed would allow a user to complete the data organization process within BioMon more efficiently.

## ***Continue contacting known sources of data***

Specifically:

- **Poll participants in courses regarding their efforts in establish plots at regular intervals after the course**
- **Periodically contact researchers who have submitted data to inquire about new plots, additional censuses, and updated contact information**
- **Maintain contact information for researchers submitting data within the metadata fields in BioMon**

Further improvement upon the plot data already at MAB necessitates the addition of data to expand the depth of its collection. This can be done by contacting established sources of data.

Researchers who have participated in biodiversity training courses offered by MAB should be contacted to determine the status of any data collection plots they intended to or have gathered data from at predetermined time intervals after the course has concluded. Contact information collected by MAB for this purpose was found to be complete, and as such, should continue to be collected and archived in the current fashion.

In addition to researchers establishing plots, those who have previously submitted plot data to MAB should be contacted on a recurring basis to request any new data they may have collected. This task should be completed, preferably, while contacting course participants so data gathering may be more focused.

Maintenance of contact information corresponding to researchers who submit data should be conducted on a regular basis. A period of no more than one year should pass between contacts to ensure contact information for plots with the correct telephone number,

mailing or email address, and organization information are up to date. These updated and corrected records should be reflected in the metadata associated with corresponding plots.

### ***Continue to make data available***

Specifically:

- **Post data summaries on MAB website, and include instructions for interpreting statistics**
- **Submit new data sets to SALVIAS**

One of the ultimate goals of the project as envisioned by MAB was diffusion of the organized plot data. To this end, all efforts towards making plot data available should be continued. The MAB website and the SALVIAS network provide valuable tools through which data can be shared.

The MAB website should be the foremost location for posting summary data, which should hold more value to the general public than raw data that they are not necessarily equipped to analyze. Summaries and instructions for interpreting such summaries should thus be posted in an organized fashion and clearly labeled according to their origins. Instructions for statistical interpretation should be posted conspicuously alongside the summary data. Contact information should be included for parties interested in additional information or viewing the entire data set.

We also recommend posting newly collected data with SALVIAS as a feasible way to reach researchers who will benefit from accessing actual raw plot data. Before the data are posted however, they should be cleaned and organized according to the steps described in the data handling manual. Continued support of the SALVIAS data collection could potentially lead to further contacts regarding additional plot data.

## 8 References

- Comiskey, James, & Dallmeier, F. (2004). *Methodology for establishing biodiversity plots*. Retrieved September 23, 2004, from <http://nationalzoo.si.edu/ConservationAndScience/MAB/researchprojects/protocols/methodology.cfm>.
- Forest Inventory and Analysis National Program Office. (2004). *Forest Inventory and Analysis*. Retrieved December 8, 2004, from <http://fia.fs.fed.us/about.htm>
- Missouri Botanical Garden. (2004). *Missouri Botanical Garden*. Retrieved October 27, 2004, from <http://www.mobot.org>
- National Zoological Park. (2004a). *Monitoring and Assessment of Biodiversity*. Retrieved September 23, 2004, from <http://nationalzoo.si.edu/ConservationandScience/MAB/>.
- National Zoological Park. (2004b). *Smithsonian National Zoological Park*. Retrieved December 8, 2004, from <http://nationalzoo.si.edu/AboutUs/Mission/>.
- O'Grady, Sally. (1999). *BioMon/win user guide*. Retrieved September 23, 2004, from <http://nationalzoo.si.edu/ConservationAndScience/MAB/biomon/BioMonManual.pdf>.
- Ramakrishnan, Raghu, & Gehrke, Johannes. (2000). *Database management systems*. Boston: McGraw-Hill.
- Smithsonian Institution. (2004). *Smithsonian Institution 2003 Annual Report*. Washington, DC: U.S. Smithsonian Books.
- Sun Microsystems. (2004). *Java Technology*. Retrieved December 8, 2004, from <http://java.sun.com>
- Thuraisingham, Bhavani M. (1997). *Data management systems: Evolution and interoperation*. Boca Raton: CRC Press.
- University of Arizona. (2004). *The SALVIAS Project*. Retrieved September 29, 2004, from <http://eeb37.biosci.arizona.edu/~salvias/main.html>

## 9 Appendices

|   |    |
|---|----|
| Appendix A: Sponsor Description .....                   | 56 |
| Appendix B: Pre-Project Plot Inventory .....            | 61 |
| Appendix C: Data Handling Manual .....                  | 62 |
| Appendix D: Summary Analyses.....                       | 63 |
| Appendix E: Previously Sent Data Letter .....           | 64 |
| Appendix F: Intent to Establish Biodiversity Plots..... | 65 |
| Appendix G: Education Course Participants .....         | 66 |
| Appendix H: Software Tool Source Code .....             | 67 |



## ***Appendix A: Sponsor Description***

Founded in 1986, at the Smithsonian Institution, the Monitoring and Assessment of Biodiversity Program (MAB) was established to “set protocol standards for the measurement and composition of vegetation in the international network of Man and the Biosphere Reserves around the world and provide biodiversity conservation training associated to the research.” (A. Alonso, personal communication, December 12, 2004) Then Smithsonian Secretary Robert M. Adams and UNESCO Director-General Amadou-Mahtar M'Bow founded MAB from UNESCO's original Man and the Biosphere Program. Up until 1994 the MAB program was under the Smithsonian's Assistant Secretary for Science, when Robert Hoffman retired. During subsequent Secretary I. Michael Heyman's administration, the Program moved to the office of Biodiversity Programs. Starting in 1999 MAB became associated with the Conservation and Research Center (CRC), a division of the National Zoo. This was the occurrence of an internal reorganization of the National Museum of Natural History. As a result, MAB became more closely associated with the CRC due to their frequent program interactions.

The Smithsonian Institution (2004, pp. 20-21) receives most of its revenue from the federal government yearly through appropriations pre-signed by Congress. The Smithsonian also receives funding through grants for specific projects. Between the fiscal years of 2002-2004 the Smithsonian received from \$1.5 billion to \$1.7 billion yearly, with 2004 marking the most money the Smithsonian has ever received to date. From the \$1.7 billion total received, the National Zoological Park receives between \$30 million to \$32 million yearly. This money is used to fund special programs, research, animal maintenance, and employee salaries.

The Smithsonian Institution's (2004, p. 38) policy is set by its secretary, appointed by the Board of Regents. The Board of Regents is made up of the Vice-President of the United States, the Chief Justice of the United States (who is the chancellor), three members of Senate, three members of the House of Representatives, and nine private citizens.

The mission of the Smithsonian (2004) is "the increase and diffusion of knowledge among men" (p. 38). The National Zoological Park (2004b, Mission) and its Conservation and Research center narrow this goal to promoting leadership and research in conservation science. The MAB program, specifically, aims to further awareness and leadership in the conservation of global biodiversity through education and research. MAB's main research goals are to promote biodiversity data collection and to increase knowledge of ecological functions so that decision-makers can reach informed conservation-based land management decisions.

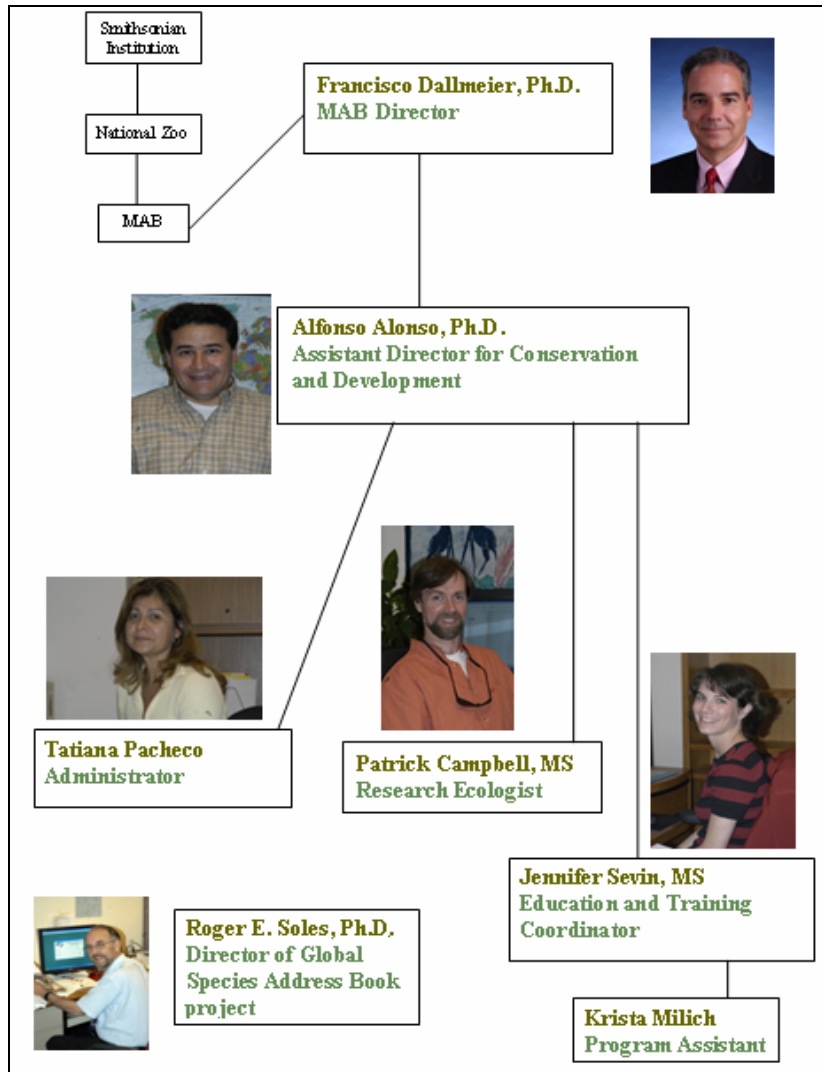
MAB's Research Program has four main objectives:

- Test and implement protocols for long-term, multi-taxa monitoring of forests.
- Establish biodiversity assessment and monitoring projects to further regional conservation needs.
- Provide data management and analytical procedures that allow rapid assessment and dissemination of information.
- Coordinate the International Biodiversity Monitoring Network (IBMN) to facilitate information exchange, information dissemination, and data quality standards formation. IBMN has a large number of research sites.

(National Zoo, 2004b)

MAB (National Zoological Park, 2004a, MAB Staff) is led by its Director, Francisco Dallmeier, who has coordinated over sixty research and training ventures worldwide and

advised many national and international committees. Other key figures in MAB's operations include Alfonso Alonso, Jennifer Sevin, and Patrick Campbell (see Figure A1 for an organizational chart). Dr. Alonso, the Director of Conservation, is involved in project planning, design of sampling protocols, and teaching for many of MAB's conferences and courses. Ms. Sevin, the Education and Training Director, coordinates MAB courses and Smithsonian Environmental Leadership courses and works to create partnerships with other organizations towards achieving MAB objectives. Mr. Campbell works on a number of biodiversity monitoring projects around the globe and also teaches courses in monitoring vegetation and mammals. These staff members coordinate their expertise to promote conservation efforts taken on by their office. They also serve as contact points for other organizations and efforts around the world concerned with biodiversity.



**Figure A1: MAB organizational chart**

MAB's main focus is the flora and vegetation of North America, Latin America, South America, and Africa (National Zoological Park, 2004a). Plant species in these regions are often considered the most valuable from a scientific viewpoint. Trees, specifically, offer a great deal of information on a locale's status to those interested in examining them. People wanting access to these collected data include managers with stewardship responsibilities, scientists and students of all ages, policymakers, educators, eco-tourists, and nature enthusiasts. Collecting data regarding the impact of human and environmental influences on

such ecologically-sensitive sites is critical in the biodiversity monitoring process, and the ability to effectively manage and distribute these data is critical to their value.

*Appendix B: Pre-Project Plot Inventory*

| Entry | Name  | Region | Country  | Plot # | Year | Re-census? | Format | Original File directory | Data found<br>[y/n?] | Species list<br>[y/n?] | Individual tree<br>[y/n?] | Notes                        |
|-------|---|--------|----------|--------|------|------------|--------|-------------------------|----------------------|------------------------|---------------------------|------------------------------|
| 1     | Dikolo  | AF     | Cameroon | 1      | 1996 |            | csv    | Backup\Dikolo           | Y                    | Y                      | Y                         |                              |
| 2     | Ejagham Forest Reserve                        | AF     | Cameroon | 4      | 1997 | Y          | biomon | Ejagham                 | Y                    | Y                      | Y                         |                              |
| 3     | Ejagham Forest Reserve                        | AF     | Cameroon | 4      |      | 2002       | biomon | Ejagham                 | Y                    | Y                      | Y                         |                              |
| 4     | Ejagham Forest Reserve                        | AF     | Cameroon | 5      | 1997 | Y          | biomon | Ejagham                 | Y                    | Y                      | Y                         |                              |
| 5     | Ejagham Forest Reserve                        | AF     | Cameroon | 5      |      | 2002       | biomon | Ejagham                 | Y                    | Y                      | Y                         |                              |
| 6     | Korup National Park                           | AF     | Cameroon | 1      | 2000 |            | biomon | korup                   | Y                    | Y                      | Y                         | Whittaker plots              |
| 7     | Korup National Park                           | AF     | Cameroon | 2      | 2000 |            | biomon | korup                   | Y                    | Y                      | Y                         | Whittaker plots              |
| 8     | Korup National Park                           | AF     | Cameroon | 3      | 2000 |            | biomon | korup                   | Y                    | Y                      | Y                         | Whittaker plots              |
| 9     | Korup National Park                           | AF     | Cameroon | 4      | 2000 |            | biomon | korup                   | Y                    | Y                      | Y                         | Whittaker plots              |
| 10    | Korup National Park                           | AF     | Cameroon | 5      | 2000 |            | biomon | korup                   | Y                    | Y                      | Y                         | Whittaker plots              |
| 11    | Korup National Park                           | AF     | Cameroon | 6      | 2000 |            | biomon | korup                   | Y                    | Y                      | Y                         | Whittaker plots              |
| 12    | Takamanda Forest Reserve, Akwaya Sub-Division | AF     | Cameroon | 6      | 2001 |            | biomon | takamanda               | Y                    | Y                      | Y                         | include Whittakers           |
| 13    | Takamanda Forest Reserve, Akwaya Sub-Division | AF     | Cameroon | 7      | 2001 |            | biomon | takamanda               | Y                    | Y                      | Y                         | include Whittakers           |
| 14    | Takamanda Forest Reserve, Akwaya Sub-Division | AF     | Cameroon | 8      | 2001 |            | biomon | takamanda               | Y                    | Y                      | Y                         | include Whittakers           |
| 15    | Takamanda Forest Reserve, Akwaya Sub-Division | AF     | Cameroon | 9      | 2001 |            | biomon | takamanda               | Y                    | Y                      | Y                         | include Whittakers           |
| 16    | Takamanda Forest Reserve, Akwaya Sub-Division | AF     | Cameroon | 10     | 2001 |            | biomon | takamanda               | Y                    | Y                      | Y                         | include Whittakers           |
| 17    | Takamanda Forest Reserve, Akwaya Sub-Division | AF     | Cameroon | 11     | 2001 |            | biomon | takamanda               | Y                    | Y                      | Y                         | include Whittakers           |
| 18    | Takamanda Forest Reserve, Akwaya Sub-Division | AF     | Cameroon | 12     | 2001 |            | biomon | takamanda               | Y                    | Y                      | Y                         | include Whittakers           |
| 19    | Takamanda Forest Reserve, Akwaya Sub-Division | AF     | Cameroon | 13     | 2001 |            | biomon | takamanda               | Y                    | Y                      | Y                         | include Whittakers           |
| 20    | Takamanda Forest Reserve, Akwaya Sub-Division | AF     | Cameroon | 14     | 2001 |            | biomon | takamanda               | Y                    | Y                      | Y                         | include Whittakers           |
| 21    | Takamanda Forest Reserve, Akwaya Sub-Division | AF     | Cameroon | 15     | 2001 |            | biomon | takamanda               | Y                    | Y                      | Y                         | include Whittakers           |
| 22    | Abakaliki                                     | AF     | Nigeria  | 1      | 1997 |            | csv    | Backup\Nigeria          | Y                    | Y                      | Y                         |                              |
| 23    | Abakaliki                                     | AF     | Nigeria  | 2      | 1997 |            | csv    | Backup\Nigeria          | Y                    | Y                      | Y                         |                              |
| 24    | Akamkpa                                       | AF     | Nigeria  | 1      | 1997 |            | csv    | Backup\Nigeria          | Y                    | Y                      | Y                         |                              |
| 25    | Akamkpa                                       | AF     | Nigeria  | 2      | 1997 |            | csv    | Backup\Nigeria          | Y                    | Y                      | Y                         |                              |
| 26    | EH  | AF     | Nigeria  | 1      | 1997 |            | csv    | Backup\Nigeria          | Y                    | Y                      | Y                         |                              |
| 27    | Okwangwo                                      | AF     | Nigeria  | 1      | 1997 |            | csv    | Backup\Nigeria          | Y                    | Y                      | Y                         |                              |
| 28    | Umukabia                                      | AF     | Nigeria  | 1      | 1997 |            | csv    | Backup\Nigeria          | Y                    | Y                      | Y                         |                              |
| 29    | Akamkpa                                       | AF     | Nigeria  | 1      | 1997 |            | biomon | nigeria                 | Y                    | Y                      | Y                         | rescode AK                   |
| 30    | Akamkpa                                       | AF     | Nigeria  | 2      | 1997 |            | biomon | nigeria                 | Y                    | Y                      | Y                         | rescode AK                   |
| 31    | Okwangwo                                      | AF     | Nigeria  | 1      | 1997 |            | biomon | nigeria                 | Y                    | Y                      | Y                         | rescode OK                   |
| 32    | Abakaliki                                     | AF     | Nigeria  | 1      | 1997 |            | biomon | nigeria                 | Y                    | Y                      | Y                         | rescode AB                   |
| 33    | Abakaliki                                     | AF     | Nigeria  | 2      | 1997 |            | biomon | nigeria                 | Y                    | Y                      | Y                         | rescode AB                   |
| 34    | Umukabia                                      | AF     | Nigeria  | 4      | 1997 |            | biomon | nigeria                 | Y                    | Y                      | Y                         | rescode UM                   |
| 35    | EH  | AF     | Nigeria  | 5      | 1998 |            | biomon | nigeria                 | Y                    | Y                      | Y                         | rescode EH                   |
| 36    | Okwangwo, Cross River State                   | AF     | Nigeria  | 2      | 2003 |            | biomon | okwangwo                | Y                    | Y                      | Y                         | species list of poor quality |
| 37    | Okwangwo, Cross River State                   | AF     | Nigeria  | 3      | 2003 |            | biomon | okwangwo                | Y                    | Y                      | Y                         | species list of poor quality |
| 38    | Okwangwo, Cross River State                   | AF     | Nigeria  | 4      | 2003 |            | biomon | okwangwo                | Y                    | Y                      | Y                         | species list of poor quality |
| 39    | Okwangwo, Cross River State                   | AF     | Nigeria  | 5      | 2003 |            | biomon | okwangwo                | Y                    | Y                      | Y                         | species list of poor quality |
| 40    | Okwangwo, Cross River State                   | AF     | Nigeria  | 6      | 2003 |            | biomon | okwangwo                | Y                    | Y                      | Y                         | species list of poor quality |
| 41    | Okwangwo, Cross River State                   | AF     | Nigeria  | 7      | 2003 |            | biomon | okwangwo                | Y                    | Y                      | Y                         | species list of poor quality |
| 42    | Okwangwo, Cross River State                   | AF     | Nigeria  | 8      | 2003 |            | biomon | okwangwo                | Y                    | Y                      | Y                         | species list of poor quality |
| 43    | Okwangwo, Cross River State                   | AF     | Nigeria  | 9      | 2003 |            | biomon | okwangwo                | Y                    | Y                      | Y                         | species list of poor quality |
| 44    | Okwangwo, Cross River State                   | AF     | Nigeria  | 10     | 2003 |            | biomon | okwangwo                | Y                    | Y                      | Y                         | species list of poor quality |
| 45    | Dinghushan                                    | AS     | China    | 1      | 1993 |            | csv    | Backup\Dinghushan       | Y                    | Y                      | Y                         |                              |
| 46    | Dinghushan                                    | AS     | China    | 2      | 1993 |            | csv    | Backup\Dinghushan       | Y                    | Y                      | Y                         |                              |
| 47    | Jiangfengling                                 | AS     | China    | 1      | 1990 | Y          | csv    | Backup\Jiangfengling    | Y                    | Y                      | Y                         |                              |
| 48    | Jiangfengling                                 | AS     | China    | 1      |      | 1993       | csv    | Backup\Jiangfengling    | Y                    | Y                      | Y                         |                              |
| 49    | Escambray                                     | CR     | Cuba     | 1      | 1994 |            | csv    | Backup\Cuba             | Y                    | Y                      | Y                         | species list of poor quality |

| Entry | Name                               | Region | Country | Plot # | Year | Re-census? | Format       | Original File directory | Data found<br>[y/n?] | Species list<br>[y/n?] | Individual tree<br>[y/n?] | Notes     |
|-------|------------------------------------|--------|---------|--------|------|------------|--------------|-------------------------|----------------------|------------------------|---------------------------|-----------|
| 50    | Grenada Dry Forest Project         | CR     | Grenada | 1      | 2004 |            | biomon       | Grenada                 | Y                    | Y                      | Y                         | Whittaker |
| 51    | Grenada Dry Forest Project         | CR     | Grenada | 2      | 2004 |            | biomon       | Grenada                 | Y                    | Y                      | Y                         | Whittaker |
| 52    | Grenada Dry Forest Project         | CR     | Grenada | 3      | 2004 |            | biomon       | Grenada                 | Y                    | Y                      | Y                         | Whittaker |
| 53    | Grenada Dry Forest Project         | CR     | Grenada | 4      | 2004 |            | biomon       | Grenada                 | Y                    | Y                      | Y                         | Whittaker |
| 54    | Grenada Dry Forest Project         | CR     | Grenada | 5      | 2004 |            | biomon       | Grenada                 | Y                    | Y                      | Y                         | Whittaker |
| 55    | Grenada Dry Forest Project         | CR     | Grenada | 6      | 2004 |            | biomon       | Grenada                 | Y                    | Y                      | Y                         | Whittaker |
| 56    | Grenada Dry Forest Project         | CR     | Grenada | 7      | 2004 |            | biomon       | Grenada                 | Y                    | Y                      | Y                         | Whittaker |
| 57    | Grenada Dry Forest Project         | CR     | Grenada | 8      | 2004 |            | biomon       | Grenada                 | Y                    | Y                      | Y                         | Whittaker |
| 58    | Grenada Dry Forest Project         | CR     | Grenada | 9      | 2004 |            | biomon       | Grenada                 | Y                    | Y                      | Y                         | Whittaker |
| 59    | Grenada Dry Forest Project         | CR     | Grenada | 10     | 2004 |            | biomon       | Grenada                 | Y                    | Y                      | Y                         | Whittaker |
| 60    | Grenada Dry Forest Project         | CR     | Grenada | 11     | 2004 |            | biomon       | Grenada                 | Y                    | Y                      | Y                         | Whittaker |
| 61    | Grenada Dry Forest Project         | CR     | Grenada | 12     | 2004 |            | biomon       | Grenada                 | Y                    | Y                      | Y                         | Whittaker |
| 62    | Grenada Dry Forest Project         | CR     | Grenada | 13     | 2004 |            | biomon       | Grenada                 | Y                    | Y                      | Y                         | Whittaker |
| 63    | Grenada Dry Forest Project         | CR     | Grenada | 14     | 2004 |            | biomon       | Grenada                 | Y                    | Y                      | Y                         | Whittaker |
| 64    | Grenada Dry Forest Project         | CR     | Grenada | 15     | 2004 |            | biomon       | Grenada                 | Y                    | Y                      | Y                         | Whittaker |
| 65    | Grenada Dry Forest Project         | CR     | Grenada | 16     | 2004 |            | biomon       | Grenada                 | Y                    | Y                      | Y                         | Whittaker |
| 66    | Grenada Dry Forest Project         | CR     | Grenada | 17     | 2004 |            | biomon       | Grenada                 | Y                    | Y                      | Y                         | Whittaker |
| 67    | Grenada Dry Forest Project         | CR     | Grenada | 18     | 2004 |            | biomon       | Grenada                 | Y                    | Y                      | Y                         | Whittaker |
| 68    | Grenada Dry Forest Project         | CR     | Grenada | 19     | 2004 |            | biomon       | Grenada                 | Y                    | Y                      | Y                         | Whittaker |
| 69    | Grenada Dry Forest Project         | CR     | Grenada | 20     | 2004 |            | biomon       | Grenada                 | Y                    | Y                      | Y                         | Whittaker |
| 70    | Puerto Rico                        | CR     | USA     | 1      | 1988 | Y          | biomon       | Bisley                  | Y                    | Y                      | Y                         |           |
| 71    | Puerto Rico                        | CR     | USA     | 1      |      |            | 1990 biomon  | Bisley                  | Y                    | Y                      | Y                         |           |
| 72    | Puerto Rico                        | CR     | USA     | 1      |      |            | 1991 biomon  | Bisley                  | Y                    | Y                      | Y                         |           |
| 73    | Puerto Rico                        | CR     | USA     | 1      |      |            | 1992 biomon  | Bisley                  | Y                    | Y                      | Y                         |           |
| 74    | Puerto Rico                        | CR     | USA     | 1      |      |            | 1994 biomon  | Bisley                  | Y                    | Y                      | Y                         |           |
| 75    | Puerto Rico                        | CR     | USA     | 1      |      |            | 1996 biomon  | Bisley                  | Y                    | Y                      | Y                         |           |
| 76    | Puerto Rico                        | CR     | USA     | 1      |      |            | 1999 biomon  | Bisley                  | Y                    | Y                      | Y                         |           |
| 77    | Virgin Islands NP, St. John        | CR     | USA     | 1      | 1992 | Y          | access       | St John                 | Y                    | Y                      | Y                         |           |
|       | Virgin Islands NP, St. John        | CR     | USA     | 1      |      |            | 1995 access  | St John                 | Y                    | Y                      | Y                         |           |
| 78    | Kejimkujik                         | NA     | Canada  | 1      | 1994 |            | csv          | Backup\Kejimkujik       | Y                    | Y                      | Y                         |           |
| 79    | Kejimkujik                         | NA     | Canada  | 2      | 1994 |            | csv          | Backup\Kejimkujik       | Y                    | Y                      | Y                         |           |
| 80    | Long Point                         | NA     | Canada  | 1      | 1995 |            | csv          | Backup\Long Point       | Y                    | Y                      | Y                         |           |
| 81    | Rocky Point, DF                    | NA     | Canada  | 1      | 1995 |            | csv          | Backup\Rocky Point      | Y                    | Y                      | Y                         |           |
| 82    | Royal Road University, Victoria BC | NA     | Canada  | 1      | 1999 |            | biomon       | Royal Roads Un          | Y                    | Y                      | Y                         |           |
| 83    | Royal Road University, Victoria BC | NA     | Canada  | 2      | 1999 |            | biomon       | Royal Roads Un          | Y                    | Y                      | Y                         |           |
| 84    | Front Royal                        | NA     | U.S.    | 1      | 1993 | Y          | csv          | Backup\Front Royal      | Y                    | Y                      | Y                         |           |
| 85    | Front Royal                        | NA     | U.S.    | 1      |      |            | 1995 csv     | Backup\Front Royal      | Y                    | Y                      | Y                         |           |
| 86    | Front Royal                        | NA     | U.S.    | 2      | 1993 |            | csv          | Backup\Front Royal      | Y                    | Y                      | Y                         |           |
| 87    | Front Royal                        | NA     | U.S.    | 3      | 1994 |            | csv          | Backup\Front Royal      | Y                    | Y                      | Y                         |           |
| 88    | Front Royal                        | NA     | U.S.    | 4      | 1994 |            | csv          | Backup\Front Royal      | Y                    | Y                      | Y                         |           |
| 89    | Front Royal                        | NA     | U.S.    | 5      | 1995 |            | csv          | Backup\Front Royal      | Y                    | Y                      | Y                         |           |
| 90    | Front Royal                        | NA     | U.S.    | 6      | 1995 |            | csv          | Backup\Front Royal      | Y                    | Y                      | Y                         |           |
| 91    | Front Royal                        | NA     | U.S.    | 8      | 1996 |            | csv          | Backup\Front Royal      | Y                    | Y                      | Y                         |           |
| 92    | Front Royal                        | NA     | U.S.    | 9      | 1997 |            | csv          | Backup\Front Royal      | Y                    | Y                      | Y                         |           |
| 93    | CRC, Front Royal, Virginia         | NA     | USA     | 1      | 1993 | Y          | biomon       | Front Royal             | Y                    | Y                      | Y                         |           |
| 94    | CRC, Front Royal, Virginia         | NA     | USA     | 1      |      |            | 1995 biomon  | Front Royal             | Y                    | Y                      | Y                         |           |
| 95    | Beni BR                            | SA     | Bolivia | 1      | 1987 | Y          | Paradox      | Bolivia                 | Y                    | Y                      | Y                         |           |
| 96    | Beni BR                            | SA     | Bolivia | 1      |      |            | 1990 Paradox | Bolivia                 | Y                    | Y                      | Y                         |           |
| 97    | Beni BR                            | SA     | Bolivia | 1      |      |            | 1991 Paradox | Bolivia                 | Y                    | Y                      | Y                         |           |



| Entry | Name                        | Region | Country  | Plot # | Year | Re-census? | Format | Original File directory | Data found<br>[y/n?] | Species list<br>[y/n?] | Individual tree<br>[y/n?] | Notes           |
|-------|-----------------------------|--------|----------|--------|------|------------|--------|-------------------------|----------------------|------------------------|---------------------------|-----------------|
| 98    | Beni BR                     | SA     | Bolivia  | 1      |      |            | 1992   | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 99    | Beni BR                     | SA     | Bolivia  | 1      |      |            | 1996   | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 100   | Beni BR                     | SA     | Bolivia  | 2      | 1987 | Y          |        | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 101   | Beni BR                     | SA     | Bolivia  | 2      |      |            | 1990   | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 102   | Beni BR                     | SA     | Bolivia  | 2      |      |            | 1991   | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 103   | Beni BR                     | SA     | Bolivia  | 2      |      |            | 1992   | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 104   | Beni BR                     | SA     | Bolivia  | 2      |      |            | 1996   | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 105   | Beni BR                     | SA     | Bolivia  | 3      | 1987 | Y          |        | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 106   | Beni BR                     | SA     | Bolivia  | 3      |      |            | 1990   | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 107   | Beni BR                     | SA     | Bolivia  | 3      |      |            | 1991   | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 108   | Beni BR                     | SA     | Bolivia  | 3      |      |            | 1992   | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 109   | Beni BR                     | SA     | Bolivia  | 3      |      |            | 1996   | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 110   | Beni BR                     | SA     | Bolivia  | 4      | 1987 | Y          |        | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 111   | Beni BR                     | SA     | Bolivia  | 4      |      |            | 1990   | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 112   | Beni BR                     | SA     | Bolivia  | 4      |      |            | 1991   | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 113   | Beni BR                     | SA     | Bolivia  | 4      |      |            | 1992   | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 114   | Beni BR                     | SA     | Bolivia  | 4      |      |            | 1996   | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 115   | Beni BR                     | SA     | Bolivia  | 5      | 1994 | Y          |        | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 116   | Beni BR                     | SA     | Bolivia  | 5      |      |            | 1996   | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 117   | Beni BR                     | SA     | Bolivia  | 6      | 1994 | Y          |        | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 118   | Beni BR                     | SA     | Bolivia  | 6      |      |            | 1996   | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 119   | Beni BR                     | SA     | Bolivia  | 7      | 1995 | Y          |        | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 120   | Beni BR                     | SA     | Bolivia  | 7      |      |            | 1997   | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 121   | Beni BR                     | SA     | Bolivia  | 8      | 1995 | Y          |        | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 122   | Beni BR                     | SA     | Bolivia  | 8      |      |            | 1997   | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 123   | Beni BR                     | SA     | Bolivia  | 9      | 1995 | Y          |        | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 124   | Beni BR                     | SA     | Bolivia  | 9      |      |            | 1997   | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 125   | Beni BR                     | SA     | Bolivia  | 10     | 1995 | Y          |        | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 126   | Beni BR                     | SA     | Bolivia  | 10     |      |            | 1997   | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 127   | Beni BR                     | SA     | Bolivia  | 11     | 1995 | Y          |        | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 128   | Beni BR                     | SA     | Bolivia  | 11     |      |            | 1997   | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 129   | Beni BR                     | SA     | Bolivia  | 12     | 1995 |            |        | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 130   | Beni BR                     | SA     | Bolivia  | 13     | 1995 | Y          |        | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 131   | Beni BR                     | SA     | Bolivia  | 13     |      |            | 1997   | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 132   | Beni BR                     | SA     | Bolivia  | 14     | 1995 | Y          |        | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 133   | Beni BR                     | SA     | Bolivia  | 14     |      |            | 1997   | Paradox Bolivia         | Y                    | Y                      | Y                         |                 |
| 134   | Choco NP                    | SA     | Colombia | 1      | 2002 |            |        | biomon Colombia         | Y                    | Y                      | N                         | no tree data    |
| 135   | Kwakwani                    | SA     | Guyana   | 3      | 1995 |            |        | Paradox Guyana          | Y                    | Y                      | N                         | no species list |
| 136   | Kwakwani                    | SA     | Guyana   | 4      | 1995 |            |        | Paradox Guyana          | Y                    | Y                      | N                         | no species list |
| 137   | Kwakwani                    | SA     | Guyana   | 5      | 1995 |            |        | Paradox Guyana          | Y                    | Y                      | N                         | no species list |
| 138   | Kwakwani                    | SA     | Guyana   | 7      | 1995 |            |        | Paradox Guyana          | Y                    | Y                      | N                         | no species list |
| 139   | Soberania                   | SA     | Panama   | 1      | 1997 |            |        | csv Backup/PN Soberania | Y                    | Y                      | Y                         |                 |
| 140   | Soberania NP                | SA     | Panama   | 1      | 1997 |            |        | biomon Soberania        | Y                    | Y                      | Y                         |                 |
| 141   | Reserva Biologica Itabo     | SA     | Paraguay | 1      | 1994 |            |        | biomon itaipu           | Y                    | Y                      | Y                         | rescode IT      |
| 142   | Reserva Biologica Limoy     | SA     | Paraguay | 1      | 1994 |            |        | biomon itaipu           | Y                    | Y                      | Y                         | rescode LI      |
| 143   | Refugio Biologico Mbaracayu | SA     | Paraguay | 1      | 1995 |            |        | biomon itaipu           | Y                    | Y                      | Y                         | rescode MB      |
| 144   | Reserva Biologica Tati Yupi | SA     | Paraguay | 1      | 1995 |            |        | biomon itaipu           | Y                    | Y                      | Y                         | rescode TY      |
| 145   | Aguas Negras                | SA     | Peru     | 1      | 1994 |            |        | csv Backup/Aguas Negras | Y                    | Y                      | Y                         |                 |
| 146   | Manu                        | SA     | Peru     | 1      | 1987 | Y          |        | csv Backup/Manu         | Y                    | N                      | Y                         |                 |

| Entry | Name       | Region | Country   | Plot # | Year | Re-census? | Format   | Original File directory | Data found<br>[y/n?] | Species list<br>[y/n?] | Individual tree<br>[y/n?] | Notes                 |
|-------|------------|--------|-----------|--------|------|------------|----------|-------------------------|----------------------|------------------------|---------------------------|-----------------------|
| 147   | Manu       | SA     | Peru      | 1      |      |            | 1991 csv | Backup\Manu             | Y                    | N                      | Y                         |                       |
| 148   | Manu       | SA     | Peru      | 2      | 1987 | Y          | csv      | Backup\Manu             | Y                    | N                      | Y                         |                       |
| 149   | Manu       | SA     | Peru      | 2      |      |            | 1991 csv | Backup\Manu             | Y                    | N                      | Y                         |                       |
| 150   | Manu       | SA     | Peru      | 3      | 1987 | Y          | csv      | Backup\Manu             | Y                    | N                      | Y                         |                       |
| 151   | Manu       | SA     | Peru      | 3      |      |            | 1991 csv | Backup\Manu             | Y                    | N                      | Y                         |                       |
| 152   | Manu       | SA     | Peru      | 4      | 1987 | Y          | csv      | Backup\Manu             | Y                    | N                      | Y                         |                       |
| 153   | Manu       | SA     | Peru      | 4      |      |            | 1991 csv | Backup\Manu             | Y                    | N                      | Y                         |                       |
| 154   | Urubamba   | SA     | Peru      | 1      | 1997 |            | access   |                         | Y                    | Y                      | Y                         |                       |
| 155   | Urubamba   | SA     | Peru      | 2      | 1997 |            | access   |                         | Y                    | Y                      | Y                         |                       |
| 156   | Urubamba   | SA     | Peru      | 3      | 1997 |            | access   |                         | Y                    | Y                      | Y                         |                       |
| 157   | Urubamba   | SA     | Peru      | 4      | 1997 |            | access   |                         | Y                    | Y                      | Y                         |                       |
| 158   | Urubamba   | SA     | Peru      | 5      | 1997 |            | access   |                         | Y                    | Y                      | Y                         |                       |
| 159   | Urubamba   | SA     | Peru      | 6      | 1997 |            | access   |                         | Y                    | Y                      | Y                         |                       |
| 160   | Urubamba   | SA     | Peru      | 7      | 1997 |            | access   |                         | Y                    | Y                      | Y                         |                       |
| 161   | Urubamba   | SA     | Peru      | 8      | 1997 |            | access   |                         | Y                    | Y                      | Y                         |                       |
| 162   | Urubamba   | SA     | Peru      | 9      | 1997 |            | access   |                         | Y                    | Y                      | Y                         |                       |
| 163   | Urubamba   | SA     | Peru      | 10     | 1997 |            | access   |                         | Y                    | Y                      | Y                         |                       |
| 164   | Urubamba   | SA     | Peru      | 11     | 1997 |            | access   |                         | Y                    | Y                      | Y                         |                       |
| 165   | Guatopo    | SA     | Venezuela | 1      | 1994 |            | csv      | Backup\Guatopo          | Y                    | Y                      | Y                         |                       |
| 166   | Los Llanos | SA     | Venezuela | 1      | 1994 |            | csv      | Backup\Los Llanos       | Y                    | Y                      | Y                         | rescodes inconsistent |
| 167   | Los Llanos | SA     | Venezuela | 2      | 1994 |            | csv      | Backup\Los Llanos       | Y                    | Y                      | Y                         | rescodes inconsistent |
| 168   | Los Llanos | SA     | Venezuela | 3      | 1994 |            | csv      | Backup\Los Llanos       | Y                    | Y                      | Y                         | rescodes inconsistent |

## *Appendix C: Data Handling Manual*



Smithsonian  
*National Zoological Park*

Monitoring and Assessment of Biodiversity  
MRC705 PO Box 37012  
Washington, DC 20013-7012  
202.633.4793 – [simab@si.edu](mailto:simab@si.edu)

[www.si.edu/simab](http://www.si.edu/simab)

# Data Handling Manual

Produced by Emmanuel Fernandez, Michael Itz, and Cory Sullivan of  
Worcester Polytechnic Institute, Worcester, MA  
December 2004

# Contents

---

|  |            |
|--|------------|
| <b>Introduction</b> .....                              | <b>iii</b> |
| <b>Quick Tips</b> .....                                | <b>iv</b>  |
| <b>Identify Data Type</b> .....                        | <b>5</b>   |
| <b>Data Conversion</b> .....                           | <b>6</b>   |
| <i>Text Files (.txt)</i> .....                         | 6          |
| <i>Access Database Files (.mdb)</i> .....              | 9          |
| <i>Excel Files (.xls)</i> .....                        | 12         |
| <i>Comma Separated Values (.csv)</i> .....             | 15         |
| <i>BioMon Files (.bak)</i> .....                       | 18         |
| <b>Start Using BioMon</b> .....                        | <b>21</b>  |
| <i>New Database</i> .....                              | 21         |
| <i>Load Site Data</i> .....                            | 22         |
| <b>Create a Regional Database</b> .....                | <b>24</b>  |
| <i>Export Species Lists</i> .....                      | 24         |
| <i>Clean and Create a Master Species List</i> .....    | 25         |
| <i>Create New Database</i> .....                       | 28         |
| <i>Import Redirected Plot Data</i> .....               | 29         |
| <i>Future Additions to the Regional Database</i> ..... | 30         |
| <b>BioMon</b> .....                                    | <b>31</b>  |
| <i>Quick Tips</i> .....                                | 31         |
| <i>Add Metadata</i> .....                              | 32         |
| <i>View Plot Data</i> .....                            | 34         |
| <b>Produce Summary Data</b> .....                      | <b>35</b>  |
| <i>Export Metadata</i> .....                           | 35         |
| <i>Run Summary Tool</i> .....                          | 36         |
| <b>Prepare for SALVIAS</b> .....                       | <b>37</b>  |
| <i>Avoid Duplication</i> .....                         | 37         |
| <i>Prepare to Send Data</i> .....                      | 38         |
| <b>Appendices</b> .....                                | <b>39</b>  |
| <i>A: Data Conversion Map</i> .....                    | 40         |
| <i>B: FAQ</i> .....                                    | 41         |

## INTRODUCTION

---

The goal of this manual is to assist in the further improvement of internal access to vegetation plot data at MAB and the availability of such data to researchers outside of MAB. Whether it be the incorporation of newly submitted data or the revision of existing data, a methodical and straightforward approach has been taken to explain the reasoning and intended outcomes of data manipulation. The following documentation can guide, from start to finish, the data handling process provided each step is followed.

The following is a generous time estimate for the initial handling of data. When one becomes more familiar with the process, significantly less time will be required.

### Start

|                          |           |
|--------------------------|-----------|
| Identify Data Type       | 5 minutes |
| Convert Format           | 10        |
| Create BioMon Database   | 5         |
| Work on Species List     | 60        |
| Update BioMon Database   | 5         |
| Produce Summary Data     | 5         |
| Prepare for Distribution | + 5       |

### Finish

95 minutes total (approximately)

## QUICK TIPS

---






While at times steps may seem unnecessary or redundant, they have been determined to be the most logical and straightforward way to go about data handling.

- ✚ It is highly recommended that a backup copy of any data you attempt to manipulate be created in a separate, clearly identified folder.
- ✚ Following each step verbatim may prevent the loss of data, yet unforeseen complications may arise and create an irreversible loss of data.
- ✚ Proper deletion of the aforementioned backup copy must be considered, so as not to confuse yourself or any future data handlers as to its purpose. Use the proper directories when working with your backup data and newly created data.
- ✚ Be aware that some steps may be unnecessary as the data you are working with may already be prepared for use with BioMon. These extra steps are used for redundancy to prevent a corrupt data import once imported into BioMon.
- ✚ There are quicker, more difficult ways to go about the same processes described here. This is a guide to easily describe every necessary step. As you become more comfortable with the following methods for data handling you can then explore alternative ways keeping in mind the proper order of events.

## IDENTIFY DATA TYPE

---

Each of the above file descriptions are ways to store organized data. However, the method by which each one goes about its organization is different. Refer to “Appendix A: Data Conversion Map” for more information. Correctly identify the type of data by either the icon or file extension and proceed to the appropriate chapter.

| Text Document   | Access Database   | Excel Spreadsheet   | Comma Separated Value   | BioMon Database File  |
|---|---|---|---|---|
|  |  |  |  |  |
| .txt  | .mdb  | .xls  | .csv  | .bak  |



## DATA CONVERSION

---

The following represent steps that have been created to improve the clarity of this manual. There are alternative methods to ones presented in this document, but they are more detailed and could result in a macro data loss if not done properly. Therefore simpler steps were devised and the reasoning behind them is listed to emphasize the importance of seemingly unnecessary steps.

- The data columns in your files may not always be labeled. Identifying which column is which may require some deductive reasoning.
- Excel tables are used to manipulate the order of data so it will import properly into BioMon
- The linear data processing steps all end with data in BioMon format and style

### Text Files (.txt)

---

#### Step 1 – Import into Excel

Open Excel:

File→Open

Change the “Files of type” drop down list to read “Text Files”

Select the appropriate file

Click “Open”

Text Import Wizard Step 1 of 3

Choose option “delimited” or “fixed width” as appropriate.

Select the appropriate row to “Start import row at” (usually 1)

“File origin” should be set to “437 : OEM United States”

Click “Next”

Text Import Wizard Step 2 of 3

Choose the appropriate “delimiters” and “text qualifier”

DO NOT select “Treat consecutive delimiters as one”

Verify these options are correct in the Data preview field

Click “Next”

Text Import Wizard Step 3 of 3

For all columns, “Column data format” is “General”

Click “Finish”

At this point, the text file will now be opened in an Excel spreadsheet. If the data did not convert correctly, or there appears to be a problem, please refer to the *Questions* section for further assistance.

---

#### Step 2 – Save plot data

File→Save As...

Select the appropriate directory  
 Select the "Save as type:" to be "CSV (Comma delimited)(\*.csv)"  
 Select the "File name" according to the following convention:

**\*\*\* Your data will not import correctly if the proper file name is not used\*\*\***

Example: "US104.csv"

US – Reserve code, only two characters  
 1 – Plot number (if greater than 9, use lower case letters, starting with "a")  
 04 – Last two numbers of census year (ex. "2004")  
 .csv – File extension

A dialogue box will open asking, "Do you want to keep the workbook in this format?"  
 Click "Yes"  
 Keep this file open

### Step 3 – Create a species list for the plot

In the ".csv" plot data file:

Copy the Species code, Family, Genus, Species, Authorship, Common name, and Specimen type columns that are present in the file  
 Paste into a new Excel spreadsheet  
 Arrange the columns in the following order:

| Species code | Family | Genus | Species | Authorship |  |  | Common name | Specimen type |
|--------------|--------|-------|---------|------------|--|--|-------------|---------------|
|--------------|--------|-------|---------|------------|--|--|-------------|---------------|

- Add any missing columns.
- If no value exists for a field, leave it blank.
- Columns F and G are "blank" placeholders for BioMon.
- For each row, if no "Specimen type" is given, insert "Tree". This will act as a placeholder so BioMon identifies that all 9 fields of information are present.

Delete all of row 1 if it reads the column headings; row 2 becomes row 1  
 Go to the last row and delete the "□" if it appears  
 Keep this species list file open

### Step 4 – Save species list

File→Save As...

Select the "Save as type:" to be "CSV (Comma delimited)(\*.csv)"  
 Select the "File name" according to the following convention:

Example: "spp-US104.csv"

spp – Identifies the file as being a species list  
 US – Reserve code  
 1 – Plot number (if greater than 9, use lower case letters, starting with "a")

04 – Last two numbers of census year (ex. “2004”)  
 .csv – File extension

A dialogue box will open asking, “Do you want to keep the workbook in this format?”  
 Click “Yes”  
 Close the species list file

---

### Step 5 – Format plot data

Return to the original “.csv” file created from the “.txt” format  
 Arrange the columns in the following order:

| Quadrat | Tree # | Stem # | X | Y | Species Code | DBH | Status | Height | Notes | Baseline | Length to pt. A | Length to pt. B |
|---------|--------|--------|---|---|--------------|-----|--------|--------|-------|----------|-----------------|-----------------|
|---------|--------|--------|---|---|--------------|-----|--------|--------|-------|----------|-----------------|-----------------|

- Add any missing columns.
- If no value exists for a field, leave it blank.
- Column D is a “blank” column as a placeholder for BioMon.
- For each row, if no “Length to point B” (column N) is given, insert “0” (zero). This will act as a placeholder so BioMon identifies that all 14 fields of information are present.

Delete any columns not listed above (ex. Family or Genus)  
 Correct the “Quadrat” column if necessary; the fields must be 3 or 4 numbers in length

Example:        125    “1” – Plot number        “25” – Quadrat number  
                   1225   “12” – Plot number        “25” – Quadrat number

Delete all of row 1 if it reads the column headings; row 2 becomes row 1  
 Go to the last row and delete the “□” if it appears  
 Save  
 Close the plot data file

---

### Step 6 – Change “.csv” to “.bak”

Open the folder containing the species list and plot data “.csv” files  
 Check to make sure file extensions are not hidden  
 Right click on one of the files  
 Select “Rename”  
 Change the file extension to “.bak”  
 A dialogue box may appear asking, “Are you sure you want to change it?”  
 Click yes  
 Repeat for the remaining files  
 Close the folder directory  
 Your files are now ready for BioMon!

## Access Database Files (.mdb)

---

### Step 1 – Create species list

Open the Access database file

Copy the Species code, Family, Genus, Species, Authorship, Common name, and Specimen type columns that are present in the file

Paste into a new Excel spreadsheet

Arrange the columns in the following order:

| Species code | Family | Genus | Species | Authorship |  |  | Common name | Specimen type |
|--------------|--------|-------|---------|------------|--|--|-------------|---------------|
|--------------|--------|-------|---------|------------|--|--|-------------|---------------|

- Add any missing columns.
- If no value exists for a field, leave it blank.
- Columns F and G are “blank” placeholders for BioMon.
- For each row, if no “Specimen type” is given, insert “Tree”. This will act as a placeholder so BioMon identifies that all 9 fields of information are present.

Delete all of row 1 if it reads the column headings; row 2 becomes row 1

Go to the last row and delete the “□” if it appears

Keep this species list file open

Keep the Access database file open

---

### Step 2 – Save species list

File→Save As...

Select the “Save as type:” to be “CSV (Comma delimited)(\*.csv)”

Select the “File name” according to the following convention:

Example: “spp-US104.csv”

spp – Identifies the file as being a species list

US – Reserve code

1 – Plot number (if greater than 9, use lower case letters, starting with “a”)

04 – Last two numbers of census year (ex. “2004”)

.csv – File extension

A dialogue box will open asking, “Do you want to keep the workbook in this format?”

Click “Yes”

Close the species file

---

### Step 3 – Create plot data file

In the Access database file:

Copy the Quadrat, Tree #, Stem #, X, Y, Species code, DBH, Status, Height, Notes, Baseline, Length to point A, and Length to point B columns

Paste the columns in a new Excel spreadsheet  
 Arrange the columns in the following order:

| Quadrat | Tree # | Stem # |  | X | Y | Species Code | DBH | Status | Height | Notes | Baseline | Length to pt. A | Length to pt. B |
|---------|--------|--------|--|---|---|--------------|-----|--------|--------|-------|----------|-----------------|-----------------|
|---------|--------|--------|--|---|---|--------------|-----|--------|--------|-------|----------|-----------------|-----------------|

- Add any missing columns.
- If no value exists for a field, leave it blank.
- Column D is a “blank” column as a placeholder for BioMon.
- For each row, if no “Length to point B” (column N) is given, insert “0” (zero). This will act as a placeholder so BioMon identifies that all 14 fields of information are present.

Correct the “Quadrat” column if necessary; the field must be 3 or 4 numbers in length

Example:        125    “1” – Plot number        “25” – Quadrat number  
                   1225   “12” – Plot number        “25” – Quadrat number

Delete all of row 1 if it reads the column headings; row 2 becomes row 1  
 Go to the last row and delete the “□ “ if it appears

#### Step 4 – Save plot data

In Excel, File→Save As...

Select the appropriate directory

Select the “Save as type:” to be “CSV (Comma delimited)(\*.csv)”

Select the “File name” according to the following convention:

**\*\*\* Your data will not import correctly if the proper file name is not used\*\*\***

Example: “US104.csv”

US – Reserve code, only two characters

1 – Plot number (if greater than 9, use lower case letters, starting with “a”)

04 – Last two numbers of census year (ex. “2004”)

.csv – File extension

A dialogue box will open asking, “Do you want to keep the workbook in this format?”

Click “Yes”

Close the plot data file

Close the Access database file

#### Step 5 – Change “.csv” to “.bak”

Open the folder containing the species list and plot data “.csv” files

Check to make sure file extensions are not hidden

Right click on one of the files

Select “Rename”

Change the file extension to “.bak”

A dialogue box may appear asking, “Are you sure you want to change it?”

Click yes

Repeat for the remaining files

Close the folder directory

Your files are now ready for BioMon!

## Excel Files (.xls)

---

### Step 1 – Save plot data

Open the Excel plot data file

Go to File→Save As...

Select the appropriate directory

Select the “Save as type:” to be “CSV (Comma delimited)(\*.csv)”

Select the “File name” according to the following convention:

**\*\*\* Your data will not import correctly if the proper file name is not used\*\*\***

Example: “US104.csv”

US – Reserve code, only two characters

1 – Plot number (if greater than 9, use lower case letters, starting with “a”)

04 – Last two numbers of census year (ex. “2004”)

.csv – File extension

A dialogue box will open asking, “Do you want to keep the workbook in this format?”

Click “Yes”

Keep this file open

---

### Step 2 – Create species list

In the “.csv” plot data file:

Copy the Species code, Family, Genus, Species, Authorship, Common name, and Specimen type columns that are present in the file

Paste into a new Excel spreadsheet

Arrange the columns in the following order:

| Species code | Family | Genus | Species | Authorship |  |  | Common name | Specimen type |
|--------------|--------|-------|---------|------------|--|--|-------------|---------------|
|--------------|--------|-------|---------|------------|--|--|-------------|---------------|

- Add any missing columns.
- If no value exists for a field, leave it blank.
- Columns F and G are “blank” placeholders for BioMon.
- For each row, if no “Specimen type” is given, insert “Tree”. This will act as a placeholder so BioMon identifies that all 9 fields of information are present.

Delete all of row 1 if it reads the column headings; row 2 becomes row 1

Go to the last row and delete the “□ “ if it appears

Keep this species list file open

---

### Step 3 – Save species list

Select the “Save as type:” to be “CSV (Comma delimited)(\*.csv)”

Select the “File name” according to the following convention:

Example: “spp-US104.csv”

spp – Identifies the file as being a species list

US – Reserve code

1 – Plot number (if greater than 9, use lower case letters, starting with “a”)

04 – Last two numbers of census year (ex. “2004”)

.csv – File extension

A dialogue box will open asking, “Do you want to keep the workbook in this format?”

Click “Yes”

Close the species list file

#### Step 4 – Format plot data

Return to the original “.csv” file created

Arrange the columns in the following order:

| Quadrat | Tree # | Stem # |  | X | Y | Species Code | DBH | Status | Height | Notes | Baseline | Length to pt. A | Length to pt. B |
|---------|--------|--------|--|---|---|--------------|-----|--------|--------|-------|----------|-----------------|-----------------|
|---------|--------|--------|--|---|---|--------------|-----|--------|--------|-------|----------|-----------------|-----------------|

- Add any missing columns.
- If no value exists for a field, leave it blank.
- Column D is a “blank” column as a placeholder for BioMon.
- For each row, if no “Length to point B” (column N) is given, insert “0” (zero). This will act as a placeholder so BioMon identifies that all 14 fields of information are present.

Delete any columns not listed above (ex. Family or Genus)

Correct the “Quadrat” column if necessary; the field must be 3 or 4 numbers in length

Example:        125    “1” – Plot number        “25” – Quadrat number  
                   1225   “12” – Plot number        “25” – Quadrat number

Delete all of row 1 if it reads the column headings; row 2 becomes row 1

Go to the last row and delete the “□” if it appears

Save

Close the species list file

#### Step 5 – Change “.csv” to “.bak”

Open the folder containing the species list and plot data “.csv” files

Check to make sure file extensions are not being hidden

Right click on one of the files

Select “Rename”

Change the file extension to “.bak”

A dialogue box may appear asking, “Are you sure you want to change it?”



Click yes  
Repeat for the remaining files  
Close the folder directory  
Your files are now ready for BioMon!

## Comma Separated Values (.csv)

---

**\*\*\*If absolutely sure data is in correct BioMon layout, skip to Step 5\*\*\***

---

### Step 1 – Format the plot data

Open the file in Excel

Arrange the columns in the following order:

| Quadrat | Tree # | Stem # |  | X | Y | Species Code | DBH | Status | Height | Notes | Baseline | Length to pt. A | Length to pt. B |
|---------|--------|--------|--|---|---|--------------|-----|--------|--------|-------|----------|-----------------|-----------------|
|---------|--------|--------|--|---|---|--------------|-----|--------|--------|-------|----------|-----------------|-----------------|

- Add any missing columns.
- If no value exists for a field, leave it blank.
- Column D is a “blank” column as a placeholder for BioMon.
- For each row, if no “Length to point B” (column N) is given, insert “0” (zero). This will act as a placeholder so BioMon identifies that all 14 fields of information are present.

Delete any columns not listed above (ex. Family or Genus)

Correct the “Quadrat” column if necessary; the field must be 3 or 4 numbers in length

Example:        125    “1” – Plot number        “25” – Quadrat number  
                   1225   “12” – Plot number        “25” – Quadrat number

Delete all of row 1 if it reads the column headings; row 2 becomes row 1

Go to the last row and delete the “□” if it appears

Keep this plot data file open

---

### Step 2 – Save plot data

File→Save As...

Select the appropriate directory

Select the “Save as type:” to be “CSV (Comma delimited)(\*.csv)”

Select the “File name” according to the following convention:

**\*\*\* Your data will not import correctly if the proper file name is not used\*\*\***

Example: “US104.csv”

US – Reserve code, only two characters

1 – Plot number (if greater than 9, use lower case letters, starting with “a”)

04 – Last two numbers of census year (ex. “2004”)

.csv – File extension

A dialogue box will open asking, “Do you want to keep the workbook in this format?”

Click “Yes”

Close the plot data file

---

### Step 3 – Format species list

Open the species list file in Excel  
 Arrange the columns in the following order:

| Species code | Family | Genus | Species | Authorship |  |  | Common name | Specimen type |
|--------------|--------|-------|---------|------------|--|--|-------------|---------------|
|--------------|--------|-------|---------|------------|--|--|-------------|---------------|

- Add any missing columns.
- If no value exists for a field, leave it blank.
- Columns F and G are “blank” placeholders for BioMon.
- For each row, if no “Specimen type” is given, insert “Tree”. This will act as a placeholder so BioMon identifies that all 9 fields of information are present.

Delete all of row 1 if it reads the column headings; row 2 becomes row 1  
 Go to the last row and delete the “□” if it appears  
 Keep this species list file open

---

### Step 4 – Save species list

Select the “Save as type:” to be “CSV (Comma delimited)(\*.csv)”  
 Select the “File name” according to the following convention:

Example: “spp-US104.csv”  
 spp – Identifies the file as being a species list  
 US – Reserve code  
 1 – Plot number (if greater than 9, use lower case letters, starting with “a”)  
 04 – Last two numbers of census year (ex. “2004”)  
 .csv – File extension

A dialogue box will open asking, “Do you want to keep the workbook in this format?”  
 Click “Yes”  
 Close the species list file

---

### Step 5 – Change “.csv” to “.bak”

Open the folder containing the species list and plot data “.csv” files  
 Check to make sure file extensions are not being hidden  
 Right click on one of the files  
 Select “Rename”  
 Change the file extension to “.bak”  
 A dialogue box may appear asking, “Are you sure you want to change it?”  
 Click yes

Repeat for the remaining files  
Close the folder directory  
Your files are now ready for BioMon!

## BioMon Files (.bak)

---

**\*\*\*The following steps are meant to check the data.  
This process may not be necessary in most cases!\*\*\***

---

### Step 1 – Change to “.csv” file format

Open the folder containing the species list and plot data “.bak” files  
 Check to make sure file extensions are not being hidden  
 Right click on one of the files  
 Select “Rename”  
 Change the file extension to “.csv”  
 A dialogue box may appear asking, “Are you sure you want to change it?”  
 Click yes  
 Repeat for the remaining files  
 Close the folder directory

---

### Step 2 – Format plot data

Open the file in Excel  
 Arrange the columns in the following order:

| Quadrat | Tree # | Stem # |  | X | Y | Species Code | DBH | Status | Height | Notes | Baseline | Length to pt. A | Length to pt. B |
|---------|--------|--------|--|---|---|--------------|-----|--------|--------|-------|----------|-----------------|-----------------|
|---------|--------|--------|--|---|---|--------------|-----|--------|--------|-------|----------|-----------------|-----------------|

- Add any missing columns.
- If no value exists for a field, leave it blank.
- Column D is a “blank” column as a placeholder for BioMon.
- For each row, if no “Length to point B” (column N) is given, insert “0” (zero). This will act as a placeholder so BioMon identifies that all 14 fields of information are present

Delete any columns not listed above (ex. Family or Genus)  
 Correct the “Quadrat” column if necessary; the field must be 3 or 4 numbers in length

Example:        125    “1” – Plot number        “25” – Quadrat number  
                   1225   “12” – Plot number        “25” – Quadrat number

Delete all of row 1 if it reads the column headings; row 2 becomes row 1  
 Go to the last row and delete the “□” if it appears  
 Keep this plot data file open

---

### Step 2 – Save plot data

File→Save As...

Select the appropriate directory

Select the "Save as type:" to be "CSV (Comma delimited)(\*.csv)"

Select the "File name" according to the following convention:

**\*\*\* Your data will not import correctly if the proper file name is not used\*\*\***

Example: "US104.csv"

US – Reserve code, only two characters

1 – Plot number (if greater than 9, use lower case letters, starting with "a")

04 – Last two numbers of census year (ex. "2004")

.csv – File extension

A dialogue box will open asking, "Do you want to keep the workbook in this format?"

Click "Yes"

Close the plot data file

#### Step 4 – Format species list

Open the species list file in Excel

Arrange the columns in the following order:

| Species code | Family | Genus | Species | Authorship |  |  | Common name | Specimen type |
|--------------|--------|-------|---------|------------|--|--|-------------|---------------|
|--------------|--------|-------|---------|------------|--|--|-------------|---------------|

- Add any missing columns.
- If no value exists for a field, leave it blank.
- Columns F and G are "blank" placeholders for BioMon.
- For each row, if no "Specimen type" is given, insert "Tree". This will act as a placeholder so BioMon identifies that all 9 fields of information are present.

Delete all of row 1 if it reads the column headings; row 2 becomes row 1

Go to the last row and delete the "□" if it appears

Keep this species list file open

#### Step 5 – Save species list

Select the "Save as type:" to be "CSV (Comma delimited)(\*.csv)"

Select the "File name" according to the following convention:

Example: "spp-US104.csv"

spp – Identifies the file as being a species list

US – Reserve code

1 – Plot number (if greater than 9, use lower case letters, starting with "a")

04 – Last two numbers of census year (ex. "2004")

.csv – File extension

A dialogue box will open asking, "Do you want to keep the workbook in this format?"  
Click "Yes"  
Close the species list file

---

**Step 6** – Change ".csv" to ".bak"

Open the folder containing the species list and plot data ".csv" files  
Check to make sure file extensions are not being hidden  
Right click on one of the files  
Select "Rename"  
Change the file extension to ".bak"  
A dialogue box may appear asking, "Are you sure you want to change it?"  
Click yes  
Repeat for the remaining files  
Close the folder directory  
Your files are now ready for BioMon!

# START USING BIOMON

---

## New Database

---

Open BioMon

Database→Create Database (Destination)

Type in the desired location of the new database folder

- Spaces are not allowed, use the underscore key “\_” if needed
- This is where BioMon will create 32 of its own files
- Files BioMon creates in this folder cannot be manipulated outside of BioMon
- If one wishes to delete the information and start over, the entire folder must be recreated (deleted then begin with **Step 1** again)

Click “OK”

Verify the correct folder has been created in the “Import/Export Target” and “Destination DB” fields on the main BioMon window



## Load Site Data

---

### Import Species List

---

#### **Step 1** – Import species list into BioMon

Open BioMon, if not already opened

Database→Open Database (Source)

Locate the folder in which the proper data is stored, already in BioMon format

Click “OK”

Verify the correct folder has been opened in the “Import/Export Target” and “Destination DB” fields on the main BioMon window

Utilities→Species Import (BioMon/DOS format)

Locate the corrected species list (“.bak”) for the desired data site

Click “Open”

---

#### **Step 2** – Verify import has worked

Trees→TreeMon

Maintenance→Species

Review the list and verify the data is correct

## Import Plot Data

---

**\*\*\*Always make sure plot data is modified so BioMon will be able to accept it\*\*\***

---

### Step 1 – Open TreeMon

Open BioMon, if not already opened  
Database→Open Database (Source)  
Locate the folder in which the proper data is stored, already in BioMon format  
Click “OK”  
Verify the correct folder has been opened in the “Import/Export Target” and “Destination DB”  
fields on the main BioMon window  
Trees→TreeMon

---

### Step 2 – Import Data

Import/Export→BioMonDOS Import  
Locate the directory in which the data file is stored  
Select the proper BioMon file (“.bak”)  
Click “Open”

---

### Step 3 – Verify import has worked

Trees→TreeMon  
Maintenance→Plots  
Review the list and verify the data is correct

## CREATE A REGIONAL DATABASE

---

The following steps include a devised method to ensure the validity of the species list. The order of steps act as a test to ensure the list is arranged in the proper order for BioMon. If it is not, BioMon will notify the user when importing the data, and will corrupt an entire database. One would need to erase the database to which the incomplete data were imported. If need be, starting over from this point makes most sense as the data would then be in the proper arrangement.

### Export Species Lists

---

*Repeat the following steps for each "source" database (those being combined into the region):*

Open BioMon

Database→Open Database (Source)

Locate the folder in which the proper data is stored, already in BioMon format

Click "OK"

Verify the correct folder has been opened in the "Import/Export Target" and "Destination DB" fields on the main BioMon window

Utilities→Table to Ascii

Click "Tables" and locate the BioMon database folder from the correct site

Select the file "species.db"

Click "Open"

Click "Files" in the "Export ASCII" window

Choose the desired destination for the new, temporary species list (".txt")

Choose the "File name" appropriately

Click "Save"

Click "Export" in the "Export ASCII" window

Repeat for each additional "source" database

## Clean and Create a Master Species List

---

### Clean Species Lists: First Pass, Automated

---

The first phase of cleaning the species lists is automated by the species cleaning tool. This tool also looks up authorships for the species it considers.

Run the species cleaning tool ("cleaner.bat" in the BioMon tools folder)

For each species list exported above:

- Click "select file to clean"

- Find the species lists exported in the step above

- Click "Open"

- Choose a filename to save the cleaned list to

- Click "Save"

- Wait while the tool cleans the list.

- The progress bar on the bottom marks its progress, and the species it has considered appear in the list box above that.

When complete, close the species cleaning tool

## Clean Species Lists: Second Pass, Manual

---

Each of the species lists produced through the previous step must now be manually reviewed. In practice, this is often the most time consuming stage of merging the data together, but it is vital to ensure consistent taxonomy.

*For each computer-reviewed list produced in the previous step:*

Open the list in Excel as a “.csv” file

Examine column G for each row

If a row has the entry “???” followed by a note:

    Examine the family, genus, and species to the left in light of the note

    If a question mark appears after one of them, this signals it was not recognized

        Correct it (and remove the question mark)

    Correct any errors found

    Look up and fill the authorship for this species into column E

    (If this is an “unknown” or a morphospecies, you may leave authorship blank)

When done, clear all values from column G (but leave the blank column present!)

Save these changes

## Create Master Species List

---

With all of the cleaned species lists and the data which you would import into BioMon, you may now proceed to create a master species list and translate your sample data to refer to it.

Before beginning this process, ensure that you have species lists and all of the sample “.bak” files for the sites you will be merging into a region.

Run the merge tool (“merger.bat” in the BioMon tools folder)

*If you wish to append to an already-created master species list*

*Click “Synch with”*

*Find the species list to append to*

*Click open*

For each site you are merging:

Click “+” on the left pane

Find the species list for that site

Click “Open”

Select the species list file

Click “+” on the right pane

Find and select **all** of the data files for this site

Click “Open” \*

Select a folder to save to

Click the “Save to” button

Choose a folder to put the merged data in

NOTE: This folder must already exist and be empty

Click “Go”

NOTE: This process might take some time

Its progress will be marked by the bar on the bottom of the window

When complete, close the window

\* You may receive a warning that the sample data contains codes which are undefined. It is up to you how to proceed at this point. If you opt to proceed anyways, samples which reference undefined codes (codes which do not appear in that site’s species list) will be ignored.

Alternatively, you may choose not to continue and do one or both of the following:

(1) adjust the lines in the sample file to reference codes in that site’s list (or “unknown”)

(2) add codes from that site’s corresponding list to the sample file

## Create New Database

---

**\*\*\*Follow the procedure above under “Start Using BioMon, New Database” to create a target database for your region.\*\*\***

---

### Step 1 – Import the species list back into BioMon

Open BioMon, if not already opened

Database→Open Database (Source)

Locate the folder in which you have created your target region database

Click “OK”

Verify the correct folder has been opened in the “Import/Export Target” and “Destination DB” fields on the main BioMon window

Utilities→Species Import (BioMon/DOS format)

Locate the merged species list (“spp-list.bak” in the output folder from the merging tool)

Click “Open”

---

### Step 2 – Verify import has worked

Trees→TreeMon

Maintenance→Species

Review the list and verify the data is correct

---

### Step 3 – Resume import of plot data

## Import Redirected Plot Data

---

### Step 1 – Open TreeMon

Open BioMon, if not already opened

Database→Open Database (Source)

Locate the folder in which the regional database was created above

Click “OK”

Verify the correct folder has been opened in the “Import/Export Target” and “Destination DB” fields on the main BioMon window

Trees→TreeMon

---

### Step 2 – Import Data

Import/Export→BioMonDOS Import

Locate the directory in which the cleaned data file is stored

Select the proper BioMon file (“.bak)

Click “Open”

---

### Step 3 – Verify import has worked

Trees→TreeMon

Maintenance→Plots

Review the list and verify the data is correct



## Future Additions to the Regional Database

---

To add additional data to a regional database, which you have already created, implement the aforementioned steps involved in creating a new database upon all of the plots to be added. Follow the italic step in the “**Create Master Species List**” stage using the species list for the regional database which already exists. In the “**Create New Database**” stage, do not create a new database but use your already completed regional database. You must, however, import the species list again. In this process, if you receive a warning about duplicate entries choose to ignore those lines. The rest of the process steps should remain the same.

# BioMON

---

## Quick Tips

---

Always make sure...

...to work in the proper directory

...data to be imported is arranged properly so BioMon can read it

In BioMon you CAN:

Produce summary results of a plot

Produce a map of trees in a plot with BioMon ("Maps2.exe")

Double click inside a text box to pick options

In BioMon you CANNOT:

Change plot data within the program

Use a "query" function

Sort data

BioMon is Y2K compliant! Just follow all the steps correctly.

## Add Metadata

---

When entering metadata, please follow the order of entry given here. Each step depends on the previous one, as the BioMon database will automatically fill in some fields.

---

### Step 1 – Open TreeMon

Open BioMon, if not already opened  
Database→Open Database (Source)  
Locate the folder in which the proper data is stored, already in BioMon format  
Click “OK”  
Verify the correct folder has been opened in the “Import/Export Target” and “Destination DB” fields on the main BioMon window  
Trees→TreeMon

---

### Step 2 – Enter reserve information

Maintenance→Reserves  
Double click on a box in the first row  
Enter all the available information in the “Reserve Edit” window  
Click “Save Changes”  
Click “Next” if there are additional reserves you wish to enter/modify  
Close “Reserve Edit” window  
Close “Browse Reserves” window

---

### Step 3 – Enter plot information

Maintenance→Plots  
Double click on a box in the first row  
Enter all the available information in the “Edit Plot” window  
NOTE: “Description” signifies the persons responsible for the plot  
“Quad Type” is always “S” unless specified otherwise  
Click “Save Changes”  
Click “Next” if there are additional plots you wish to enter/modify  
Close “Edit Plot” window  
Close “Browse Plots” window

---

### Step 4 – Enter organization and person information

Maintenance→Organizations  
Double click in a box in the first row

Click "Clear" in the "Edit Organizations" window

Enter all the available information

NOTE: "OrgCode" can only be up to 3 letters

Double click in the "Contact Person" field

Double click in a box in the first row in the "Browse Responsible Persons" window

Enter all the available information

NOTE: "Person Code" is usually the person's initials (up to 3 letters)

Click "Save Changes"

Close "Edit Responsible Persons" window

Close "Browse Responsible Persons" window

Click "Save Changes" in the "Edit Organizations" window

Click "Next" if there are additional organizations you wish to enter/modify

Close "Edit Organizations" window

Close "Browse Organizations" window

---

### Step 5 – Enter census information

Maintenance→Census

Double click on the first "Person Code" box (on the far right)

Enter all the available information in the "Edit Census" window

Click "Save Changes"

Click "Next" if there are additional censuses you wish to edit/modify

Close "Edit Census" window

Close "Browse Census" window

## View Plot Data

---

### Step 1 – Open TreeMon

Open BioMon, if not already opened

Database→Open Database (Source)

Locate the folder in which the proper data is stored, already in BioMon format

Click “OK”

Verify the correct folder has been opened in the “Import/Export Target” and “Destination DB” fields on the main BioMon window

Trees→TreeMon

---

### Step 2 – View data

Maintenance→Reserves

⋮

Organizations

Choose any of the menu options to view the related data

## PRODUCE SUMMARY DATA

---

### Export Metadata

---

Open BioMon

Database→Open Database (Source)

Locate the folder in which the regional database is stored, already with metadata

Click “OK”

Verify the correct folder has been opened in the “Import/Export Target” and “Destination DB” fields on the main BioMon window

Utilities→Table to Ascii

Click “Tables” and locate the BioMon database folder from the correct site

Select the file “person.db”

Click “Open”

Click “Files” in the “Export ASCII” window

Choose the desired destination for the exported file (“.txt”)

Choose the “File name” appropriately

Click “Save”

Click “Export” in the “Export ASCII” window

Repeat for each additional “source” database  
by opening the report (“.txt”)

Repeat this process for “census.db”, “reserve.db”, “organize.db”, “plot.db”

## Run Summary Tool

---

Run the summary tool ("summary.bat" in the BioMon tools folder)

Click "Set persons" and select the file you exported from "person.db" in the previous step

Repeat this procedure for "Set organizations", "Set species list" (use the list produced by the merge tool), "Set reserves", "Set plots", and "Set censuses"

Click "Load metadata"

Click "+"

Find and select (shift-click) all of the plot data from the region (produced by the merge tool)

Click "Open"

Click "Go"

Select a new, empty folder (which must already exist) to save the resulting PDF files

Click "Save"

This process may take a few moments, its progress will be marked by the bar to the left of "Go"

## PREPARE FOR SALVIAS

---

### Avoid Duplication

---

Go to the SALVIAS website <<http://www.salvias.net/pages/index.html>>

Data→Login to SALVIAS database

Type in the correct user name and password

Click “Login”

Click “Plots”

Click on the drop down box next to “Project”

Verify there will be no duplication by looking down the list



## **Prepare to Send Data**

---

Create an organized folder in which all BioMon data files for a specific site are contained. Include any pertinent metadata necessary for SALVIAS not already accounted for in the BioMon files.

# APPENDICES

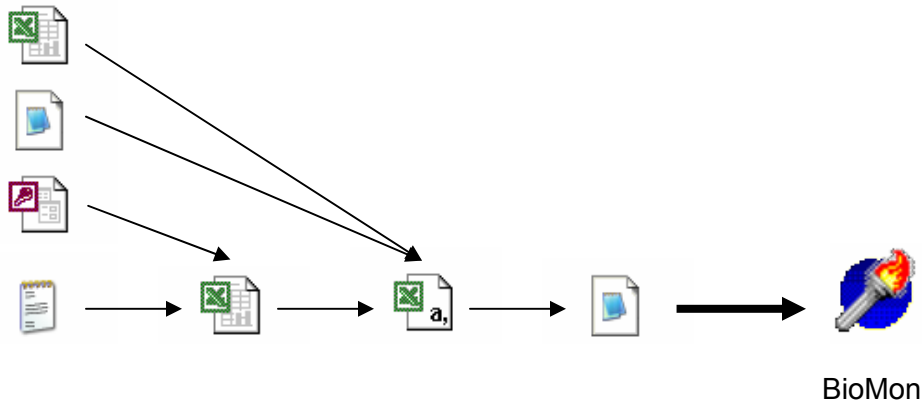
---






A: Data Conversion Map.....40  
B: FAQ .....41

# A: Data Conversion Map

---

Original Data Format      Conversion      New Data Format



| Text Document   | Access Database   | Excel Spreadsheet   | Comma Separated Value   | BioMon Database File  |
|---|---|---|---|---|
|  |  |  |  |  |
| .txt  | .mdb  | .xls  | .csv  | .bak  |

## **B: FAQ**

---

### ***What should I do if the data are separated by spaces or columns in text (.txt) format?***

Open Excel

Go to Open→File, select the file

In the Text Import Wizard Step 2 of 3, choose the “Delimiters” to be “Tab” and any others if appropriate

### ***Is it okay if the text in my comma separated values does not have quotation marks?***

Yes. This will not harm the functionality of the file. Conversion to a “.csv” file will accommodate this situation.

### ***My “.csv” file will not convert to a “.bak” file!***

Check to make sure you can view file extensions in the appropriate folder window:

Select Tools→Folder Options

Choose the “View” tab

In the “Advanced settings” window under “Files and Folders” uncheck the box marked “Hide extensions for known file types”

Click “OK”

### ***My site has more than 9 plots! What should I do about the naming convention?***

This situation may arise, in which case continue using the same naming convention, however, replace the number identifying the plot with the letter “a” in place of the desired “10”. Continue with “b” for “11”, “c” for “12”, and so on. This concept is similar to counting in hexadecimal, where letters of the alphabet are used in place of double-digit numbers.

### ***Error - “Network initialization failed.”***

This indicates BioMon was not installed correctly. Contact your system administrator for further assistance.

### ***In BioMon I don’t see the entire field of data; it appears to be cut off.***

It is truncated; BioMon has a limit on how many characters it can accept, depending on what category the data is organized under. For example: the authorship field accepts up to 20 characters, while the ResCode field only accepts up to 8. The data still exists, however. To view it, find the original file from which the data was imported into BioMon. Change the file extension to “.csv” and open it.

***“Convert all measurements to centimeters?”***

Yes.

***“Only x/7, x/14 fields present” when importing data, what should I do?***

The plot data was not properly organized when working with it in “.csv”. First, find the directory containing the BioMon files this data was entering and delete it. (This is the folder with the 32 files BioMon created when you created the BioMon database) Second, you will need to create the database, import the species list, and import any previously imported data all over again.

Smithsonian Institution  
SI/MAB Biodiversity Program  
MRC705 PO Box 37012  
Washington, DC 20013-7012  
202.633.4793 - simab@si.edu

[www.si.edu/simab](http://www.si.edu/simab)



Smithsonian  
*National Zoological Park*

## ***Appendix D: Summary Analyses***

The following is an example of summary analyses from Kejimikujik, Canada created with the summary production tool developed by the project team.

**PLOT SUMMARY INFORMATION FOR**

**Kejimkujik**  
Nova Scotia, Canada

reserve code KJ  
located at 44.36 N 65.26 W

**FOR STEMS 10cm AND LARGER  
BY FAMILY AND SPECIES**

**CONTENTS**

|               |             |
|---------------|-------------|
| <b>PLOT 1</b> | <b>1994</b> |
| <b>PLOT 2</b> | <b>1994</b> |

*Smithsonian Institution*  
Monitoring and Assessment of Biodiversity Program  
MRC705 PO Box 37012  
Washington, DC 20013-7012  
202.633.4793 – simab@si.edu



**Census** – whether this is the first time this plot was examined or an additional time

**Longitude and latitude** – the approximate location of this plot

**Basal Area** – the sum of the cross-sectional areas at breast height for trees of a family/species

**Number of Species** – the number of species in the plot of a family

**Plot count** – how this plot is numbered

**Abundance** – the number of trees of a family/species within the plot

**Frequency** – the number of subplots in which a species appears

**Contact Name** – person or persons responsible for collecting the data

**Date surveyed** – the year this data was collected

**Plot type** – may be hectare plots (1-ha), Whittaker plots, or Gentry transects

**Reserve code** – the two letter code used to identify this reserve in BioMon

**Relative Basal Area** – the fraction of the cross-sectional areas at breast height of all trees occupied by a family/species

**Family Importance Value (FIV)** – the sum of relative abundance, relative basal area, and relative number of species; used for ranking importance of families within a plot

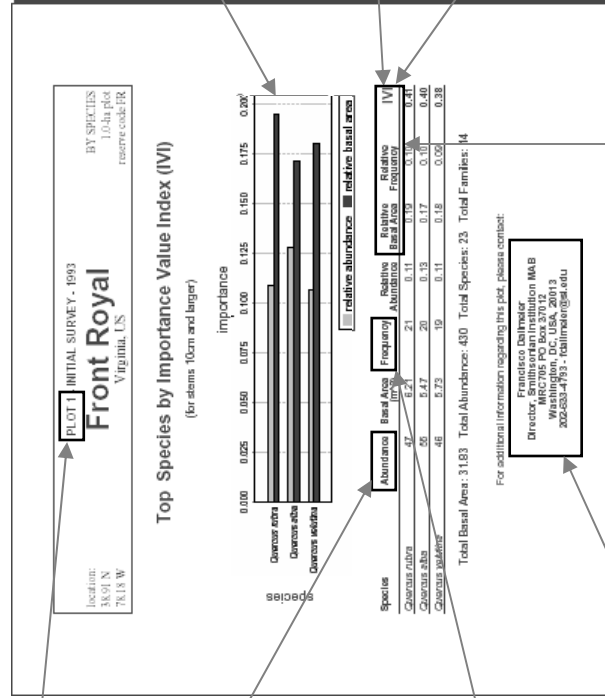
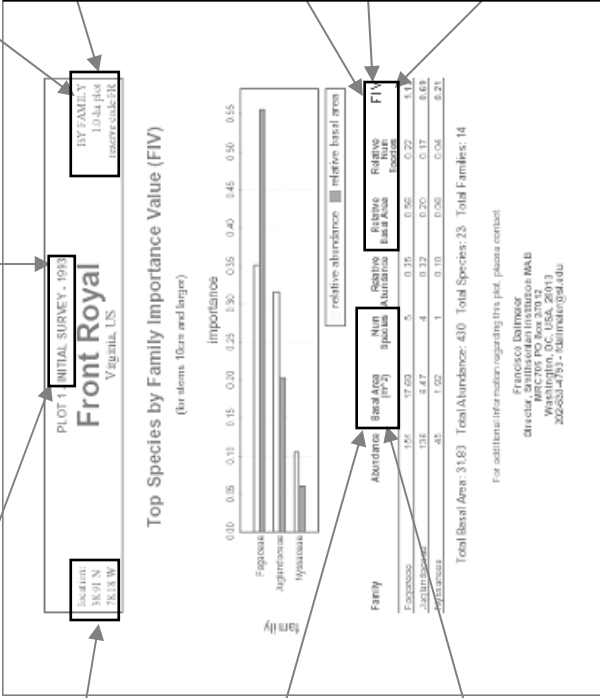
**Relative Number of Species** – the fraction of species in a plot belonging to a family

**Graph** – shows relative basal area and relative abundance for visual comparison, with family/species ranked by FIV/IVI

**Relative Frequency** – the fraction of the frequency of a species divided by the sum of all species' frequencies

**Relative Abundance** – the fraction of all trees of a family/species within the plot

**Importance Value Index (IVI)** – the sum of relative abundance, relative basal area, and relative frequency; used for ranking importance of species within a plot



PLOT 1 - INITIAL SURVEY - 1994

location:  
44.36 N  
65.26 W

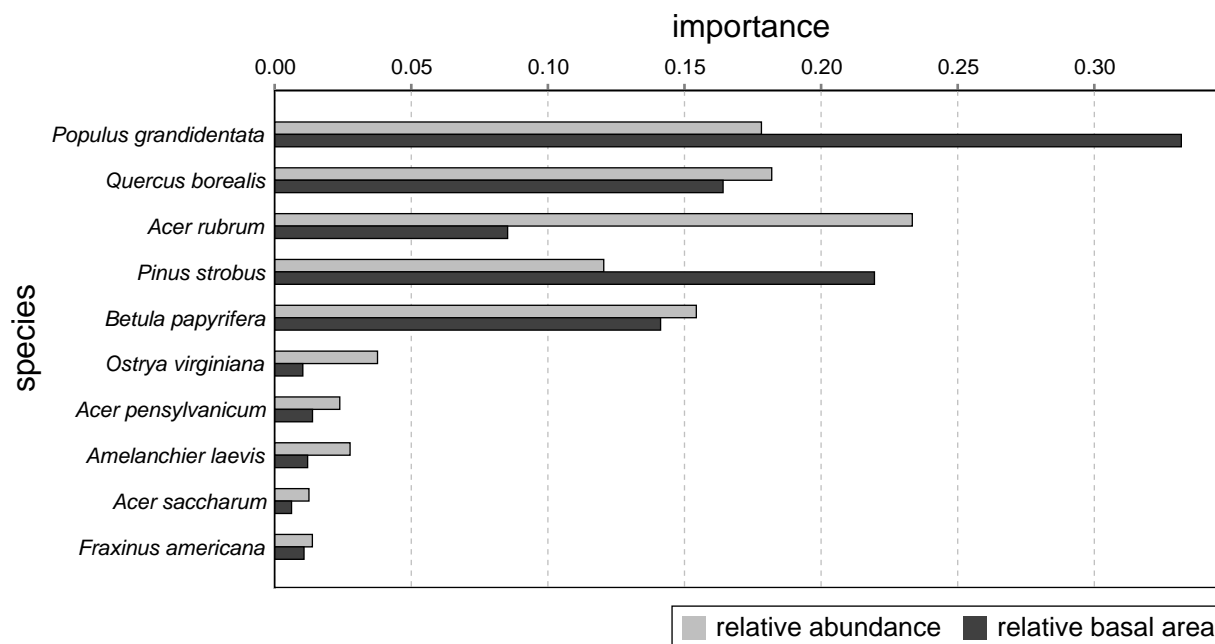
# Kejimikujik

Nova Scotia, Canada

BY SPECIES  
1.0-ha plot  
reserve code KJ

## Top Ten Species by Importance Value Index (IVI)

(for stems 10cm and larger)



| Species                      | Abundance | Basal Area (m <sup>2</sup> ) | Frequency | Relative Abundance | Relative Basal Area | Relative Frequency | IVI         |
|------------------------------|-----------|------------------------------|-----------|--------------------|---------------------|--------------------|-------------|
| <i>Populus grandidentata</i> | 142       | 10.73                        | 18        | 0.18               | 0.33                | 0.11               | <b>0.62</b> |
| <i>Quercus borealis</i>      | 145       | 5.31                         | 23        | 0.18               | 0.16                | 0.14               | <b>0.48</b> |
| <i>Acer rubrum</i>           | 186       | 2.76                         | 25        | 0.23               | 0.09                | 0.15               | <b>0.47</b> |
| <i>Pinus strobus</i>         | 96        | 7.10                         | 21        | 0.12               | 0.22                | 0.13               | <b>0.47</b> |
| <i>Betula papyrifera</i>     | 123       | 4.57                         | 21        | 0.15               | 0.14                | 0.13               | <b>0.42</b> |
| <i>Ostrya virginiana</i>     | 30        | 0.33                         | 13        | 0.04               | 0.01                | 0.08               | <b>0.13</b> |
| <i>Acer pensylvanicum</i>    | 19        | 0.45                         | 13        | 0.02               | 0.01                | 0.08               | <b>0.12</b> |
| <i>Amelanchier laevis</i>    | 22        | 0.39                         | 12        | 0.03               | 0.01                | 0.07               | <b>0.11</b> |
| <i>Acer saccharum</i>        | 10        | 0.20                         | 8         | 0.01               | 0.01                | 0.05               | <b>0.07</b> |
| <i>Fraxinus americana</i>    | 11        | 0.35                         | 6         | 0.01               | 0.01                | 0.04               | <b>0.06</b> |

Total Basal Area: **32.32** Total Abundance: **797** Total Species: **12** Total Families: **8**

For additional information regarding this plot, please contact:

**Brian Craig**  
Network Science Advisor, Environment Canada EMAN  
867 Lakeshore Road  
Burlington, Ontario, Canada, L7R 4A6  
905-336-4431 - brian.craig@ec.gc.ca

PLOT 1 - INITIAL SURVEY - 1994

location:  
44.36 N  
65.26 W

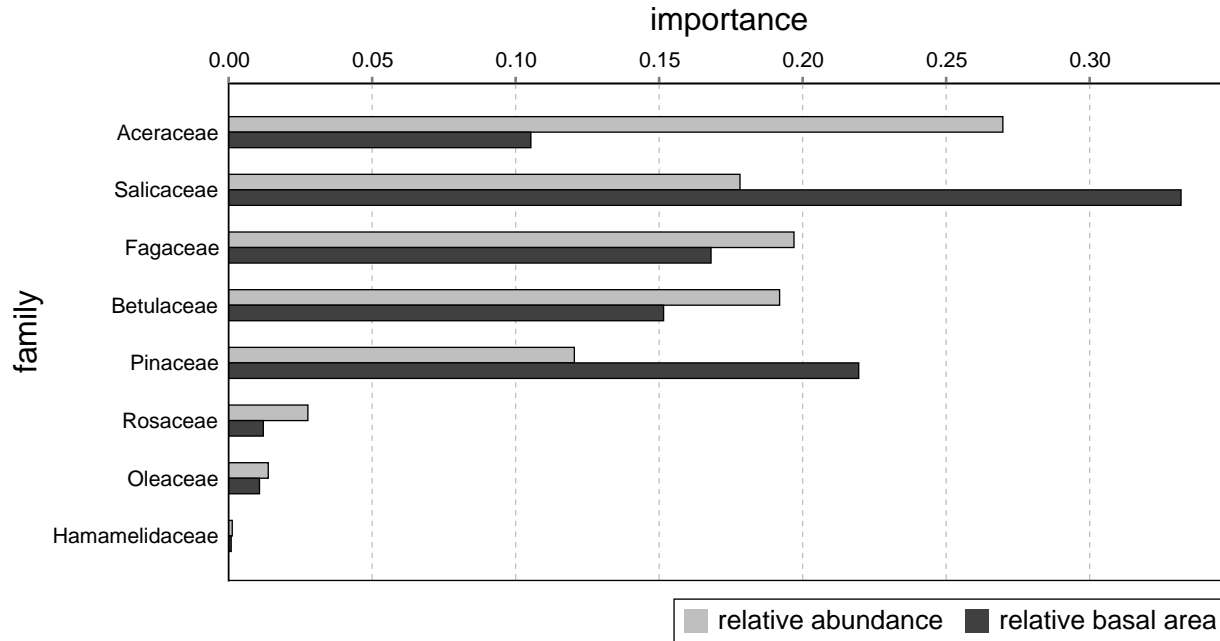
# Kejimikujik

Nova Scotia, Canada

BY FAMILY  
1.0-ha plot  
reserve code KJ

## Top Ten Species by Family Importance Value (FIV)

(for stems 10cm and larger)



| Family         | Abundance | Basal Area (m <sup>2</sup> ) | Num Species | Relative Abundance | Relative Basal Area | Relative Num Species | FIV         |
|----------------|-----------|------------------------------|-------------|--------------------|---------------------|----------------------|-------------|
| Aceraceae      | 215       | 3.40                         | 3           | 0.27               | 0.11                | 0.25                 | <b>0.63</b> |
| Salicaceae     | 142       | 10.73                        | 1           | 0.18               | 0.33                | 0.08                 | <b>0.59</b> |
| Fagaceae       | 157       | 5.43                         | 2           | 0.20               | 0.17                | 0.17                 | <b>0.53</b> |
| Betulaceae     | 153       | 4.90                         | 2           | 0.19               | 0.15                | 0.17                 | <b>0.51</b> |
| Pinaceae       | 96        | 7.10                         | 1           | 0.12               | 0.22                | 0.08                 | <b>0.42</b> |
| Rosaceae       | 22        | 0.39                         | 1           | 0.03               | 0.01                | 0.08                 | <b>0.12</b> |
| Oleaceae       | 11        | 0.35                         | 1           | 0.01               | 0.01                | 0.08                 | <b>0.11</b> |
| Hamamelidaceae | 1         | 0.03                         | 1           | 0.00               | 0.00                | 0.08                 | <b>0.09</b> |

Total Basal Area: **32.32** Total Abundance: **797** Total Species: **12** Total Families: **8**

For additional information regarding this plot, please contact:

**Brian Craig**  
Network Science Advisor, Environment Canada EMAN  
867 Lakeshore Road  
Burlington, Ontario, Canada, L7R 4A6  
905-336-4431 - brian.craig@ec.gc.ca

PLOT 2 - INITIAL SURVEY - 1994

location:  
44.36 N  
65.26 W

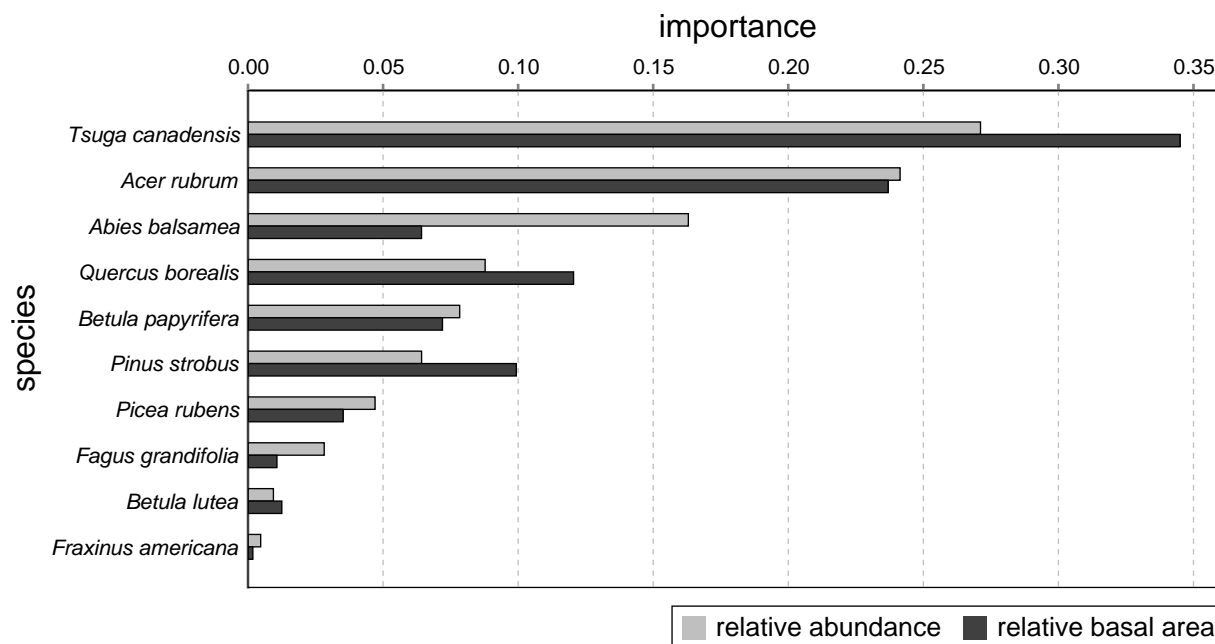
# Kejimkujik

Nova Scotia, Canada

BY SPECIES  
1.0-ha plot  
reserve code KJ

## Top Ten Species by Importance Value Index (IVI)

(for stems 10cm and larger)



| Species                   | Abundance | Basal Area (m <sup>2</sup> ) | Frequency | Relative Abundance | Relative Basal Area | Relative Frequency | IVI         |
|---------------------------|-----------|------------------------------|-----------|--------------------|---------------------|--------------------|-------------|
| <i>Tsuga canadensis</i>   | 173       | 13.48                        | 22        | 0.27               | 0.35                | 0.14               | <b>0.76</b> |
| <i>Acer rubrum</i>        | 154       | 9.26                         | 25        | 0.24               | 0.24                | 0.16               | <b>0.64</b> |
| <i>Abies balsamea</i>     | 104       | 2.51                         | 21        | 0.16               | 0.06                | 0.14               | <b>0.37</b> |
| <i>Quercus borealis</i>   | 56        | 4.71                         | 23        | 0.09               | 0.12                | 0.15               | <b>0.36</b> |
| <i>Betula papyrifera</i>  | 50        | 2.81                         | 17        | 0.08               | 0.07                | 0.11               | <b>0.26</b> |
| <i>Pinus strobus</i>      | 41        | 3.88                         | 11        | 0.06               | 0.10                | 0.07               | <b>0.24</b> |
| <i>Picea rubens</i>       | 30        | 1.38                         | 14        | 0.05               | 0.04                | 0.09               | <b>0.17</b> |
| <i>Fagus grandifolia</i>  | 18        | 0.42                         | 10        | 0.03               | 0.01                | 0.07               | <b>0.10</b> |
| <i>Betula lutea</i>       | 6         | 0.49                         | 4         | 0.01               | 0.01                | 0.03               | <b>0.05</b> |
| <i>Fraxinus americana</i> | 3         | 0.07                         | 2         | 0.00               | 0.00                | 0.01               | <b>0.02</b> |

Total Basal Area: **39.07** Total Abundance: **638** Total Species: **12** Total Families: **5**

For additional information regarding this plot, please contact:

**Brian Craig**  
Network Science Advisor, Environment Canada EMAN  
867 Lakeshore Road  
Burlington, Ontario, Canada, L7R 4A6  
905-336-4431 - brian.craig@ec.gc.ca

PLOT 2 - INITIAL SURVEY - 1994

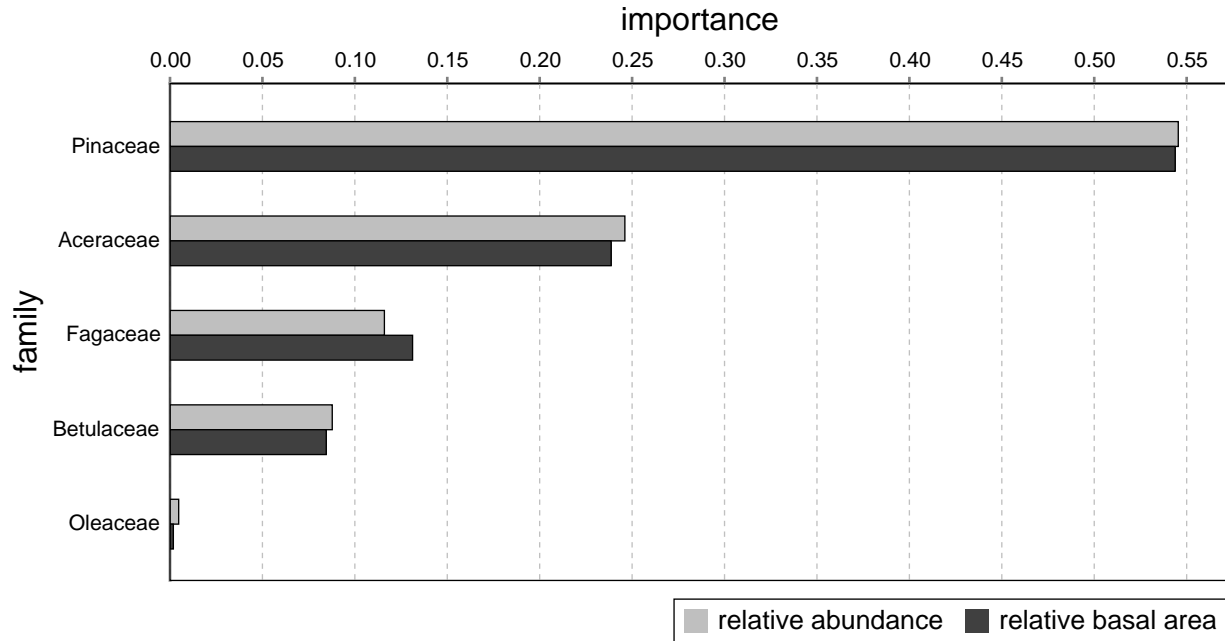
location:  
44.36 N  
65.26 W

**Kejimikujik**  
Nova Scotia, Canada

BY FAMILY  
1.0-ha plot  
reserve code KJ

**Top Ten Species by Family Importance Value (FIV)**

(for stems 10cm and larger)



| Family     | Abundance | Basal Area (m <sup>2</sup> ) | Num Species | Relative Abundance | Relative Basal Area | Relative Num Species | FIV         |
|------------|-----------|------------------------------|-------------|--------------------|---------------------|----------------------|-------------|
| Pinaceae   | 348       | 21.25                        | 4           | 0.55               | 0.54                | 0.33                 | <b>1.42</b> |
| Aceraceae  | 157       | 9.32                         | 3           | 0.25               | 0.24                | 0.25                 | <b>0.73</b> |
| Fagaceae   | 74        | 5.13                         | 2           | 0.12               | 0.13                | 0.17                 | <b>0.41</b> |
| Betulaceae | 56        | 3.30                         | 2           | 0.09               | 0.08                | 0.17                 | <b>0.34</b> |
| Oleaceae   | 3         | 0.07                         | 1           | 0.00               | 0.00                | 0.08                 | <b>0.09</b> |

Total Basal Area: **39.07** Total Abundance: **638** Total Species: **12** Total Families: **5**

For additional information regarding this plot, please contact:

**Brian Craig**  
Network Science Advisor, Environment Canada EMAN  
867 Lakeshore Road  
Burlington, Ontario, Canada, L7R 4A6  
905-336-4431 - brian.craig@ec.gc.ca

## ***Appendix E: Previously Sent Data Letter***

Dear <Contact Person>,

Please receive sincere greetings from the Smithsonian Institution/Monitoring and Assessment of Biodiversity Program. It has been a long time since Francisco and I heard from you. As you probably know, Jim is now with the National Park Service and here at MAB we continue our work on the conservation of biodiversity. We appreciate your continued partnership with our Program.

MAB is currently working with a group of college interns to update our global network of plots. We currently have data from censuses you conducted < in location and year> in our database. We would like to inquire if you have conducted any re-censuses in that region or have implemented other plots. If so, we would like to invite you to include your recent data in our global network, currently with more than 400 plots around the world. One of our objectives is to summarize all data and place it on our website for use by scientists, resource managers, educators and other interested persons wanting to learn more about biodiversity plots. We would also like to make raw data available on a controlled access website for use by other botanists and ecologist, currently managed by the University of Arizona under the SALVIAS program (<http://www.salvias.net/pages/index.html>). Credit and authorship will be given to the original data collector if data is used and published.

Please let us know your thoughts, preferably before November 26, 2004. We can certainly talk about this initiative in more detail if you are interested.

All the best,

Alfonso

Alfonso Alonso, Ph. D  
aalonso@si.edu  
202-633-4780

cc Jennifer Sevin.

## ***Appendix F: Intent to Establish Biodiversity Plots***

Dear <Contact Person>,

Greetings from the Smithsonian Institution/Monitoring and Assessment of Biodiversity Program. It has been a long time since Francisco and I heard from you, possibly since you took the biodiversity course. As you probably know, Jim is now with the National Park Service and here at MAB we continue our work on the conservation of biodiversity. We appreciate your continued partnership with our Program.

MAB is currently working with a group of college interns to update our global network of plots. Our records indicate that you established a biodiversity plot < in location and year>, but the data was not transferred to the network. We would like to inquire if you could update the status of the plot (e.g. if you conducted any re-censuses or have implemented other plots) as we invite you to include your data in our global network, currently with more than 400 plots around the world. One of our objectives is to summarize all data and place it on the internet for use by scientists, resource managers, educators and other interested persons wanting to learn more about biodiversity plots. We would also like to make raw data available on a controlled access website for use by other botanists and ecologists, currently managed by the University of Arizona under the SALVIAS program (<http://www.salvias.net/pages/index.html>). Credit and authorship will be given to the original data collector if data is used and published.

Please let us know your thoughts, preferably before November 26, 2004. We can certainly talk about this initiative in more detail if you are interested.

All the best,

Alfonso

Alfonso Alonso, Ph. D  
aalonso@si.edu  
202-633-4780

cc Jennifer Sevin.

## ***Appendix G: Education Course Participants***

Dear <Course Participant>,

Greetings from the Smithsonian Institution/Monitoring and Assessment of Biodiversity Program. I am the new Education and Training Coordinator for the MAB Program. I understand from our records that you participated in one of our courses and thus I would like to ask for your assistance.

Francisco has many new ideas for our education program that includes expanding our curricula of professional training courses and documenting the success of the training. In addition, MAB is currently documenting and expanding the biodiversity plot network and the use of Biomon. We currently have a network of more than 400 sites around the world, and we are in the process of making it available to scientists, resource managers, educators and other interested persons wanting to learn more about biodiversity plots.

You will find a questionnaire below. Please take a few moments to answer and email back your thoughts. We value very much your input. If possible, we ask that you respond to this email by November 29, 2004.

If you have any questions, please do not hesitate to contact me.

Thank you in advance for your assistance-

Sincerely,

Jennifer Sevin, MAB

-----  
Questionnaire:

- 1) Please provide us with your most current contact and job information (i.e. address, phone number, employer, job title, current projects, etc.)
- 2) Please provide us with three (or more) potential course topics that you would be interested in attending.
- 3) What do you see as a need/gap in biodiversity education and training?
- 4) What MAB course did you take? What year did you take the course?
- 5) How did the course benefit you in your profession?
- 6) Did you establish a biodiversity 1 ha plot (or a monitoring program) as a result of the MAB course? Please describe in detail.
- 7) Any additional comments or suggestions



## ***Appendix H: Software Tools and Other Electronic Resources***

*(see attached compact disc)*

### Contents:

#### ***biomon/***

Includes the four regional biomon databases: Africa, Asia, North America, and South America. Also includes originals for data sets not included in the regions.

***unmerged/*** - includes data sets which were not merged into regional databases due to species list problems

***whittaker/*** - includes data sets which have Whittaker plots that were not merged

#### ***csv/***

Contains the original comma-separated value files for each of the regional databases. Also includes exported metadata tables in \*.txt form.

#### ***summaries/***

Contains PDF summaries generated for each of the four regional databases.

#### ***BioMon32.zip***

The BioMon installation files to install BioMon on your local computer.

#### ***BioMonTools.exe***

A self-extracting archive of the software tools (merger, cleaner, summary) created to augment BioMon. Extract to c:\biomon\tools after installing BioMon.

#### ***data\_handling\_manual.pdf***

The Data Handling Manual, describing procedures for organizing and manipulating data, including use of the software tools.

#### ***iqp.pdf***

The final IQP report from the Biodiversity Database Management project.

#### ***censuses.xls***

The inventory spreadsheet generated in the first phase of the project.

#### ***slides.ppt***

The slides accompanying the final presentation delivered for the project.