Tropical Ecology, Assessment and Monitoring Network
Data Management Protocol

Version 2.0

Table of Contents

1.0. Purpose and Audience

The Data Management Protocol for the Tropical Ecology, Assessment and Monitoring (TEAM) Network describes the guidelines, processes and available hardware and software solutions for the collection, management and distribution of TEAM Network data. It does not describe the entire TEAM Information Management System (IMS) or provide technical documentation for components of the TEAM IMS. The intended audience is all TEAM Network Members involved in the collection, management or distribution of data and metadata from the TEAM Monitoring Protocols. This is a dynamic document that will be updated as the TEAM Network grows and the Information Management System evolves.

2.0. Introduction

The TEAM Network is committed to making data from all Monitoring Protocols globally accessible to the scientific and conservation communities and to the general public in as close to real time as is technically feasible. Rapid dissemination of TEAM data to the global scientific and conservation communities is crucial to maximize the utility of TEAM data for change detection and for informing the development of sound conservation strategies.

3.0. Data Management

Data management is the process of collecting, organizing, archiving and preserving data in perpetuity. The TEAM Data Management Protocol is designed to clearly explain to TEAM Network members how to utilize the TEAM Network portal and other software tools to store, archive and maintain all aspects of the data collected in-accordance with the TEAM Monitoring Protocols.

Data and Metadata

TEAM data include the field observations, collections, laboratory analysis, and post-processing of data generated from the TEAM Monitoring Protocols. To fully understand, use and preserve the integrity of these data, additional information must be collected. This is known as Metadata. Metadata are data about data. Both metadata and data are collected in a standardized process during each TEAM Monitoring Protocol Sampling Period.

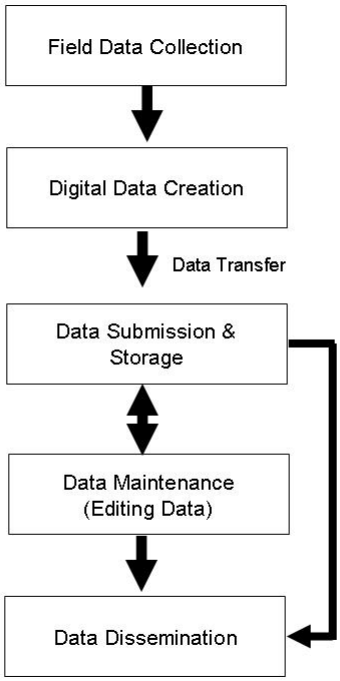
4.0. Flow of information in TEAM

The flow of information in TEAM is illustrated in Figure 1. This figure and the following sections describe in detail the steps that occur as information flows from the field or the laboratory into the Information Management System.

- Field Data Collection: Data are collected on Field Forms according to the TEAM Monitoring Protocols. All Field Forms need to be archived at the TEAM Site.

- Generating Digital Data: The Field Forms are transcribed into a digital format using the standardized Data Entry File. Data Dictionaries specify the types and ranges of the values that can be entered in each cell of the Data Entry File.
- Data Transfer, Submission and Storage: After entering the data and metadata from each Sampling Period into the Data Entry File, the file is given a standardized name and then it is saved. Data files are uploaded to the TEAM Network portal.
- Data Maintenance: All data maintenance (e.g. modifications and edits) will be done through the Data Management section of the TEAM Network portal. The complete history (i.e. all edits and modifications) will be stored in the database.
- Data Dissemination: All TEAM data are then distributed via the WWW.

Figure 1. Data Management Workflow for TEAM Monitoring Protocols.



4.1 Field Data Collection

Each time you collect TEAM data in the field or the laboratory, you should record the observations or other information on the standardized Field Forms associated with each TEAM Monitoring Protocol. Using these forms ensures consistency among different observers and different sites. These forms can be downloaded from the TEAM Network portal (www.teamnetwork.org) and printed onto waterproof paper for field use. You should check the web site each month to make sure you are always using the latest version of the Field Forms. Any time a change occurs to the protocol or field form you will also receive an email notifying you of this change.

Using the Field Forms simplifies the process of recording the data and also limits the number of data entry mistakes. In all cases, work with a partner when collecting data in the field or the lab. For example, if one person makes the observation or measurement, the second person can record the information on the Field Forms and repeat the value out loud for verification. This helps in catching mistakes during data collection, and improves the quality of the data. At the end of each day review your Field Forms to identify any obvious mistakes. This will reduce further data quality problems before entering data into the Data Entry Forms.

Spatial Data Collection

The deployment and geo-referencing of the sampling units in the field is a crucial component of every TEAM protocol. All spatial information should be collected according to the Sampling Unit Placement protocol available at www.teamnetwork.org. The Sampling Unit Placement protocol provides in explicit detail all of the procedures to be administered in and out of the field, particularly with regard to the management of all collected spatial data (e.g. initial sampling design files, GPS files, submitting geo-referenced sampling units for approval, etc).

Archive Field Forms

The paper Field Forms are a critical part of the data management process. DO NOT lose them. They are the primary source of documentation for each observation and measurement. Be sure to archive all Field Forms in a secure place and keep an inventory of the archived forms. Each TEAM Monitoring Protocol specifies the precise archival format for the Field Forms but general guidelines are described here.

An "Original" and "Back-up" Field Forms archive are required for each TEAM Monitoring Protocol. These two archives should be situated in separate geographic locations to ensure that the archives are not destroyed in case of fire or flood. Document these locations, as well as the person responsible for maintaining the archives, and send this information to the TEAM Network office. The "Back-up" archive should be updated every quarter with photocopies of the original Field Forms from data collection events occurring during that quarter.

Store all Field Forms in binders in the designated archive. Field Forms should be grouped by Sampling Period. Example Sampling Periods for the current TEAM Monitoring Protocols are monthly for climate protocol and annually for the vegetation protocol. A 2006 Climate Field Form notebook for example, will contain all the Field Forms for the year partitioned by the monthly Sampling Period. Lastly, keep a back-up copy of all Field Forms in a separate location.

Important Points

- **Designate geographic locations for the "Original" and "Back-up" Field Forms Archives.**
- **Inform the TEAM Network office of the locations and the person responsible.**

- **Setup the binders to contain the TEAM Monitoring Protocol data for one or more years.**
- **After each Sampling Period, archive the Field Forms in the "Original" Archive.**
- **Each quarter, make a photocopy of the original Field Forms for that quarter and put these copies in the Field Forms "Back-up" archive.**
-

EcoPDA

The TEAM Network is developing mobile technology with smart phones and PDAs (EcoPDAs) that can be used to gather scientific data in the field. These devices will greatly facilitate data collection while simultaneously ensuring the data are of the highest scientific quality. The first version of the EcoPDA is currently being field tested. TEAM Sites interested in field testing the EcoPDA should contact team-webmaster@teamnetwork.org . The Data Management Protocol will be updated when the EcoPDA is officially adopted for use. See http://www.teamnetwork.org/en/content/ecopda for more details on the EcoPDA.

4.2 Digital Data Creation

Once the data from a particular TEAM Monitoring Protocol are documented in the Field Forms, it should be entered **as soon as possible** into the Data Entry File that is part of the TEAM Monitoring Protocol. Ideally, the data should be entered into the Data Entry File each day, but if that is not feasible; this should be done as soon as possible, at least on a weekly basis. Each Data Entry File has a Data Dictionary, which is a table that describes all the variables in the Data Entry File, which type they are (i.e. an integer number, text, a date), the value ranges that are permitted, and examples of each. When entering data into the Data Entry File, make sure it follows the specifications outlined in the appropriate Data Dictionary. The Data Dictionaries for all the TEAM Monitoring Protocols are referenced in Appendix 5.3 and are available online at www.teamnetwork.org

Make sure that you transfer all of the information from the Field Forms to the Data Entry File. The Minimum Data Standards (Appendix 5.1) provide guidelines for reviewing the data in the Data Entry File to make sure they are complete and contain allowable values. If a species is unidentified for a record at the moment of data entry, make sure that an appropriate and informative morpho-species name is included for that record/specimen. Detailed taxonomic instructions are specified in each protocol and the corresponding Data Dictionary.

Save your data frequently when you enter information into the Data Entry File. When you save your file, use the following naming convention: Two code letter for protocol - two code letter for site – sampling period_v1.xls. The "v1" at the end of the file name means version 1. If several versions are needed increase the version numbers by one. So, for example the name "**CX-VG-2009.01_v1.xls**" means that this is the vegetation protocol file (VG), from Caxiuanã (CX), where the data were collected during the first sampling period of 2009.

At this point, the file is ready to be uploaded to the TEAM Network portal at www.teamnetwork.org. Once you have saved your file, but before you upload it, check the file thoroughly for mistakes (use the FILTER option in EXCEL to look at the range of values for each variable).

DeskTEAM

DeskTEAM is a custom software application that will generate Data Entry Files automatically. It will also assist in the transfer and upload of data to the TEAM Network portal. DeskTEAM is currently being field tested and will be released in late 2009 or early 2010. The Data Management Protocol will be updated when DeskTEAM is released for use in the TEAM Network. Contact team-webmaster@teamnetwork.org for more information on DeskTEAM.

4.3 Data Upload and Storage

- Once the Data Entry Files are complete and the files have been saved, they are ready to be uploaded to the TEAM Network portal. At each TEAM Site, and for each TEAM Standard Monitoring Protocol, it is best if there is a single designated person who is responsible for uploading the Data Entry File. The Data Management section of the TEAM Network portal is available at www.teamnetwork.org and navigating to "Products →Data→ Data Management". Instructions on how to use the Data Management section of the TEAM Network portal are available here: http://www.teamnetwork.org/en/help-data-manage . Figures 2 through 5 demonstrate how to upload a Data Entry File, fix errors and save the data to the database.

Figure 2: Upload Data by selecting appropriate information and uploading the file.

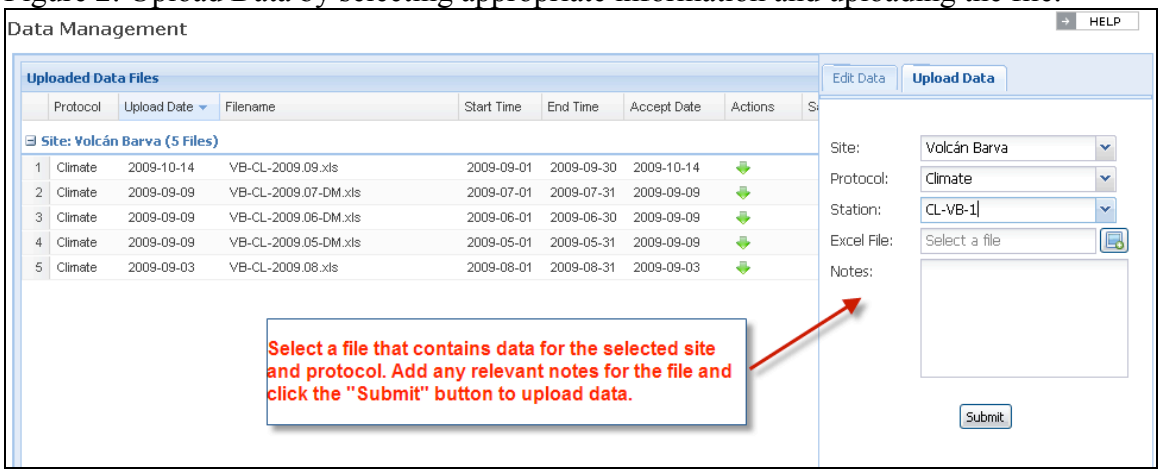


Figure 3: Press the Find Error button to find and correct errors in the data entry file.

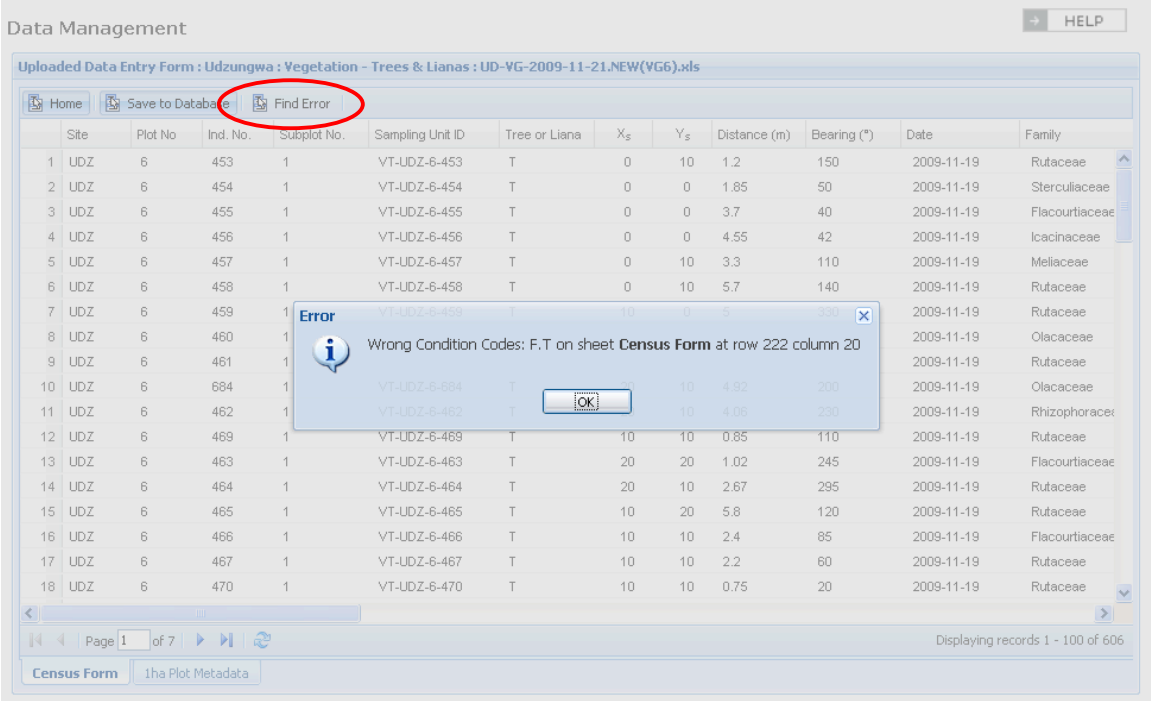


Figure 4: Repeat until No Errors are found and then push the "Save to Database" button.

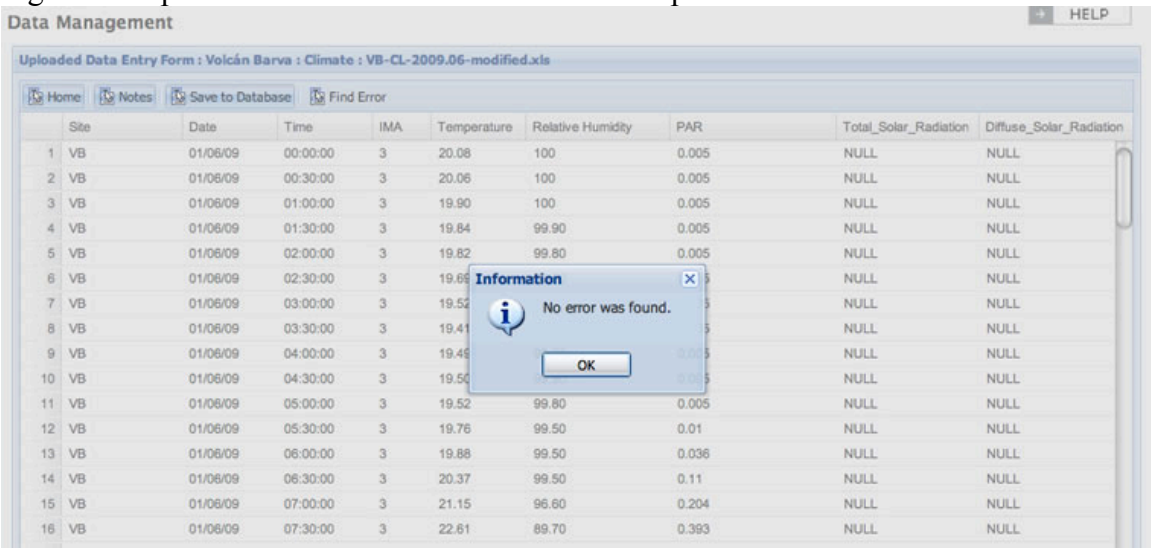
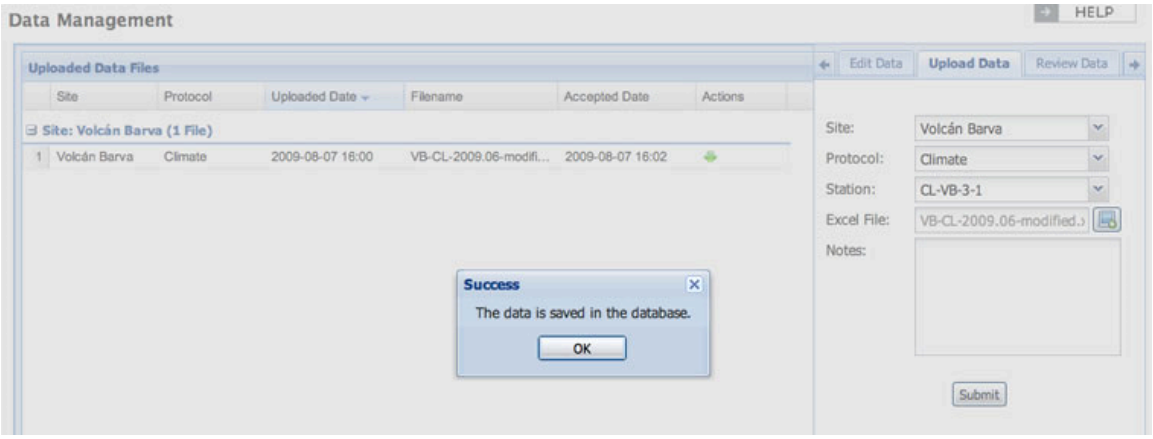


Figure 5: The Data Entry File is now saved to the database. An email notification will be sent out to confirm the Data Entry File was saved.



Important Steps

- **Upload the Data Entry File in the Data Management section of the portal.**
- **Use the "Find Error" tool to find errors in the Data Entry File. If there are too many or you want to start over you can delete the file.**
- **Once all the errors are corrected click the "Save to Database" button. Once this happen the data are saved into the system and an email will be sent out confirming that the data are saved. This ends the data upload process. Note: without saving the data to the database the process is not complete.**

Spatial Data Upload

Spatial data is also uploaded via the Data Management area of the TEAM Network Portal. All spatial data must be uploaded and approved prior to any TEAM Monitoring Protocol data uploading. This ensures that all field measurements and observations are correctly geo-referenced. Please see the Sampling Unit Placement protocol for detailed information (www.teamnetwork.org).

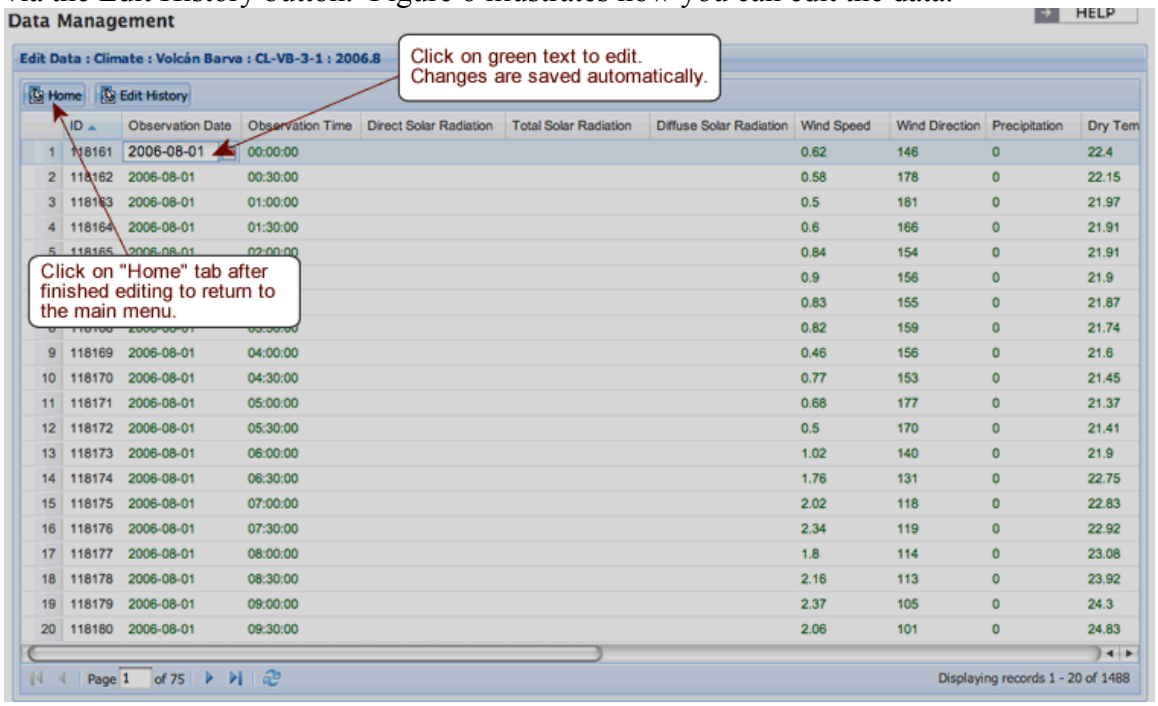
Data Entry Files can only be saved to the system if they meet the Minimum Data Standards (See Appendix 5.1).

4.4 Data Maintenance

Once the data are saved to the system there will inevitably need to be some modifications to the data. All data modifications are done via the "Edit Data" tab located in the Data Management area of the TEAM Network portal. This is available at www.teamnetwork.org by navigating to "Products →Data→ Data Management". The data have edit restrictions on them to ensure that only people with the appropriate expertise can edit the data. For example, not everyone can edit spatial information, nor can everyone at a TEAM Site edit the taxonomic information. All edit restrictions are established via the Site Management Tool that the Site Manager will setup for each site.

It is important to point out that all modifications are required to meet the same minimum data quality standards (Section 5.1) as the data being uploaded. Furthermore, every change or modification to a data record is stored. This means that if a mistake is made we

can "roll-back" the system to find the previous value. Viewing the data history is done via the Edit History button. Figure 6 illustrates how you can edit the data.



	ID	Observation Date	Observation Time	Direct Solar Radiation	Total Solar Radiation	Diffuse Solar Radiation	Wind Speed	Wind Direction	Precipitation	Dry Tem
1	118161	2006-08-01	00:00:00				0.62	146	0	22.4
2	118162	2006-08-01	00:30:00				0.58	178	0	22.15
3	118163	2006-08-01	01:00:00				0.5	181	0	21.97
4	118164	2006-08-01	01:30:00				0.6	166	0	21.91
5	118165	2006-08-01	02:00:00				0.84	154	0	21.91
							0.9	156	0	21.9
							0.83	155	0	21.87
							0.82	159	0	21.74
9	118169	2006-08-01	04:00:00				0.46	156	0	21.6
10	118170	2006-08-01	04:30:00				0.77	153	0	21.45
11	118171	2006-08-01	05:00:00				0.68	177	0	21.37
12	118172	2006-08-01	05:30:00				0.5	170	0	21.41
13	118173	2006-08-01	06:00:00				1.02	140	0	21.9
14	118174	2006-08-01	06:30:00				1.76	131	0	22.75
15	118175	2006-08-01	07:00:00				2.02	118	0	22.83
16	118176	2006-08-01	07:30:00				2.34	119	0	22.92
17	118177	2006-08-01	08:00:00				1.8	114	0	23.08
18	118178	2006-08-01	08:30:00				2.16	113	0	23.92
19	118179	2006-08-01	09:00:00				2.37	105	0	24.3
20	118180	2006-08-01	09:30:00				2.06	101	0	24.83

Important Points

- **All data edits and modifications are done via the Data Management section of the TEAM Network portal. Select the "Edit Data" tab and the appropriate TEAM Site and Protocol.**
- **All data edits will be recorded and can be "rolled-back" if needed.**
- **Make all edits and modifications as soon as possible to ensure they do not get lost or forgotten.**
- **All definitions and rules described in the Data Dictionary will be applied to all data edits and modifications.**

4.5 Data Dissemination

TEAM Monitoring Protocol data will become available for dissemination to public end-users in as near-real time as possible at http://www.teamnetwork.org/en/data/query. TEAM data will be freely distributed to anyone once end users submit their first and last name, their organization and email addresses and accept the TEAM Network Data Use Policy (Section 5.3). All the desired data will be downloaded in a Data Package. The Data Package includes the data files (in one of several file types), a metadata document that explains in detail the data file, and a README file. In the near future many processed data products as well as reports and summaries will be available from the TEAM Network portal.

4.6 Visualization, Analytical Tools and External Data Integration

The TEAM Network is committed to providing cutting-edge visualization and analytical tools for information generated from the network. Custom Google Maps applications and graphing packages are used throughout the Data Query and Download Application. Current tools are available on www.teamnetwork.org via the Data→ Tools menu. The TEAM Network is interested in integrating with information systems from all relevant organizations. Please contact team-webmaster@teamnetwork.org for more information.

4.7 Glossary

Sampling Period – A Sampling Period is period of time during which data are collected at the frequency of sampling for a TEAM Monitoring Protocol. Examples are monthly for butterflies and annually for trees and lianas. The Sampling Period time period will be different for each TEAM Monitoring Protocol.

Data Dictionary – The Data Dictionaries define the values that go into the Data Entry Files. They provide definitions, lists of acceptable values and identify if a field can be left blank and much more information.

Data Entry File – The Data Entry Files are used to transcribe the Field Forms into a digital format. The values should conform to the definitions in the DATA Dictionaries. Once complete the Data Entry Files are uploaded to TEAM Servers.

Data Package – A Data Package is a zip file that contains the data files, metadata and README documents that a portal user has decided to download.

Field Forms - Each Standardized TEAM Monitoring Protocol has one or more Field Forms. These can be printed out prior to going into the field and are the original source of observations and measurements collected using TEAM Monitoring Protocols.

TEAM Network Portal – The portal systems used by the TEAM Network to manage all information and data related to the protocols. The portal contains many software applications, databases and custom programming. Please contact team-webmaster@teamnetwork.org for technical details

5.0. APPENDICIES

5.1. Minimum Data Standards

Scientific data associated with the TEAM Monitoring Protocols must meet the following minimum data standards:

Data must conform to the current version of the Data Entry File or Web template for the relevant Protocol.

The data types, data ranges and values must comply with the specifications in the TEAM Monitoring Protocol Data Dictionaries.

Records should contain valid values for all variables that do not include NULL among its possible values, including:

All records must contain valid date and time information.

All records must contain species names or information regarding morpho-species.

All records must contain valid descriptors for spatial location.

The data file must be named according to the convention described in the Data Management Protocol.

Data must be uploaded to the TEAM Network portal at www.teamnetwork.org.

5.2. Tropical Ecology, Assessment and Monitoring (TEAM) Network Data Use Policy

The Tropical Ecology, Assessment and Monitoring (TEAM) Network Data Sharing and Use Policy outlines the TEAM Network standards for data sharing, access, authorship, citation, and restrictions of TEAM data, and applies to all contributors and users of TEAM data.

The TEAM Network of Conservation International (CI) funds and coordinates the systematic monitoring or biodiversity through a network of tropical field stations, to quantify and forecast changes in biodiversity at local, regional and global scales. TEAM aims to understand both the underlying dynamics of biodiversity, and the responses of biodiversity to major drivers of change, particularly changes in climate and land use/land cover. The TEAM Network members recognize that achieving this goal requires the coordination of an integrated and systematic sampling program at multiple spatial and temporal scales. Further, to maximize the utility of TEAM data for change detection and for informing the development of sound conservation strategies, rapid dissemination of TEAM data to the global scientific and conservation communities is crucial. Thus, the TEAM Network is committed to making TEAM data globally accessible to the scientific and conservation communities and to the general public.

TEAM Network Data Access Requirements

The access to all TEAM data is subject to requirements set forth by this policy document to enable data providers to track usage, evaluate its impact in the community, and confirm users' acceptance of the terms of acceptable use. These requirements are standardized across the TEAM Network to provide contractual exchange of data among Network Data Providers and Data Users that can be encoded into electronic form and exchanged between computers. This will allow direct access to data through a common portal once these requirements have been fulfilled. The following information is required directly or by proxy prior to the transference of any data object:

Registration:

1. Name
2. Affiliation
3. Email Address
4. Full Contact Information
5. Acceptance of the Data Use Agreement
6. A Statement of Intended Use that is compliant with the above agreements. Such statements may be submitted explicitly or made implicitly via the data access portal interface.

General Data Use Agreement

IMPORTANT: PLEASE READ CAREFULLY. Use of the Data Set (defined below) is subject to the terms and conditions of the agreement set forth below. By clicking "I AGREE," the Data User (defined below) agrees to be bound by the terms and conditions of this General Data Use Agreement ("Agreement"). If the Data User does not agree, click "**I DO NOT AGREE**." By refusing to agree to the terms provided herein, the Data User is not authorized to use the Data Set.

Definitions

"**Data Set**" – Digital data and its metadata provided through the data access portal interface, including data and metadata derived from TEAM monitoring protocols, including field observations, collections, laboratory analysis, or the post-processing of existing data and identified by a unique identifier issued by the TEAM Network.

"**Data User**" - individual to whom access may be granted to this Data Set in dependence on that Data User's acceptance of this Agreement, including his or her immediate collaboration sphere, defined here as the institutions, partners, students and staff with whom the Data User collaborates, and with whom access must be granted, in order to fulfill the Data User's intended use of the Data Set.

"**Data Set Creator**" individual or institution that produced the Data Set.

"**Data Set Distributor**" - individual or institution providing access to the Data Sets.

"**Data Set Contact**" - party designated in the accompanying metadata of the Data Set as the primary contact for the Data Set.

Conditions of Use

The re-use of scientific data has the potential to greatly increase communication, collaboration and synthesis within and among disciplines, and to improve the scientific foundations for conservation, and thus is fostered, supported and encouraged by the TEAM Network. The Data Set Creator hereby grants to the Data User a non exclusive royalty-free license to use this Data Set, subject to the following terms:

1) *Acceptable use*. Use of the Data Set is restricted to academic, research, educational, government, recreational, or other not-for-profit professional purposes. The Data User is permitted to produce and distribute derived works from this Data Set provided that those derivatives are released under the same license terms as those accompanying this Data Set. Any other uses for the Data Set or its derived products will require explicit permission from the Data Set Creator.

2) *Redistribution*. The Data Set is provided for use by the Data User. The metadata and this license must accompany all copies made and be available to all users of this Data Set. The Data User will not redistribute the original Data Set beyond this collaboration sphere.

3) *Authorship*. The Data Set has been licensed in the spirit of open scientific collaboration. The Data User agrees to notify the Data Set Creator prior to use of the data, and to provide an explanation of how the Data Set is intended to be used. In addition, the Data User agrees to offer co-authorship to the Data Set Creator if such coauthorship is requested by the Data Set Creator.

4) *Citation*. The TEAM Data are made immediately available as close as possible after field collection, however, taxonomic identification and other quality control processes may require several months to complete. Therefore, the TEAM data may undergo periodic revision and it is necessary to track Data Set versions in any derived products. Thus, the Data User agrees to properly cite the Data Set, including the Data Set Identifier, in any publications or in the metadata of any derived data products that are produced using the Data Set. Citation shall take the following general form: Creator, Year of Data Publication, Title of Dataset, Dataset identifier.

5) *Acknowledgment*. The Data User agrees to include the following acknowledgment in any publications where the Data Set contributed significantly to its content:

> "Data were provided by the TEAM Network of Conservation International, funded by the Gordon and Betty Moore Foundation."

In addition, the Data User agrees to include any additional acknowledgment of institutional support or specific funding awards provided in the metadata accompanying this Data Set in any publications where the Data Set contributes significantly to its content.

6) *Notification*. The Data User will register the citations to all publications and derivative works based on or derived from the Data Set at www.teamnetwork.org or, if the registry is not available, by sending an email message containing the complete citation to team-operations@teamnetwork.org. In addition, the Data User will notify the Data Set Contact when any derivative work or publication based on or derived from the Data Set is distributed. The Data User will provide the TEAM Network Office and the Data Contact with two reprints of any publications resulting from use of the Data Set and will provide copies, or on-line access to, any derived digital products.

By accepting this Data Set, the Data User agrees to abide by the terms of this agreement. The Data Creator and the TEAM Network shall have the right to terminate this agreement immediately by written notice upon the Data User's breach of, or noncompliance with, any of its terms. The Data User may be held responsible for any misuse that is caused or encouraged by the Data User's failure to abide by the terms of this agreement.

Disclaimer

COMPLETE ACCURACY OF DATA AND METADATA ARE NOT GUARANTEED. ALL DATA AND METADATA ARE MADE AVAILABLE "AS IS" WITHOUT WARRANTY OF ANY KIND, INCLUDING THE IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. THE DATA USER HOLDS ALL PARTIES INVOLVED IN THE PRODUCTION OR DISTRIBUTION OF THE DATA SET HARMLESS FROM DAMAGES RESULTING FROM ITS USE OR INTERPRETATION INCLUDING ANY CONSEQUENTIAL DAMAGES, SPECIAL DAMAGES OR LOST PROFITS.

Arbitration

It is the policy of the TEAM Network to make every reasonable effort to resolve all issues or disputes that may arise under this Agreement fairly by negotiation without litigation, if practicable. Any dispute arising out of or relating to this Agreement which is not settled by agreement of the parties shall be finally settled by arbitration in accordance with the UNCITRAL Arbitration Rules as at present in force. Any disputes that cannot be resolved by negotiation shall be subject to arbitration using a single arbitrator. The arbitration shall take place in Washington, DC, and the results of which shall be final, non-appealable, binding on each party, and enforceable in any court of competent jurisdiction. The terms and conditions of this Agreement shall be construed in accordance with the laws of the District of Columbia without regard to any conflicts of laws principles.

DATA USER ACKNOWLEDGES BY CLICKING “I AGREE” THAT HE/SHE HAS READ AND UNDERSTANDS THIS AGREEMENT AND AGREES TO BE BOUND BY ITS TERMS. DATA USER FURTHER AGREES THAT THIS AGREEMENT SUPERSEDES ANY PREVIOUS LICENSE, AGREEMENT, OR PROPOSAL, WHETHER WRITTEN OR ORAL, AND ANY OTHER COMMUNICATIONS RELATING TO THE SUBJECT MATTER OF THIS AGREEMENT.

5.3. Data Dictionaries

Data Dictionaries for all current and "under development" TEAM Monitoring Protocols are available at: http://www.teamnetwork.org/protocols. Navigate to the protocol of interest and download the Data Entry Form Data Dictionary.

5.4. Field Forms

Field Forms for all current and "under development" TEAM Monitoring Protocols are available at: http://www.teamnetwork.org/protocols. Navigate to the protocol of interest and download the Field Form.