Long Database Report

ForestPlots.net – managing permanent plot information across the tropics

Gabriela Lopez-Gonzalez, Mark Burkitt, Simon L. Lewis & Oliver L. Phillips

Abstract: Compiling and analysing tropical forests plot data has provided key insights into the population and carbon dynamics of tropical forests. ForestPlots.net is a web application developed to provide a secure online environment for long-term forest plot data for researchers worldwide working within international networks such as RAINFOR, AFRI TRON, and TROBIT, allowing scientists to manage, analyse, and compare their data to other sites. The underlying database (GIVD ID 00-00-001) in ForestPlots.net is a relational database which utilizes more than 50 tables to store plot location, individual taxonomic information and repeated diameter measurements for trees. Currently the database holds information on more than 800 plots from 27 countries with approximately half a million tropical trees tagged, measured, and monitored through time. The web application allows users, depending on their permission level, to view, edit, upload and download data of the plots they have access to. A novel feature of the database is the query library which produces outputs for the selected plots on biomass, basal area, wood productivity, and stem dynamics.

Keywords: carbon dynamics; long-term forest plot; relational database; tropical forest.

Abbreviations: AfriTRON = African Tropical Rainforest Observation Network; POM = point of measurement; RAINFOR = The Amazon Forest Inventory Network; TROBIT= Tropical Biomes in Transition.

Received: 1 July 2010 – Accepted: 12 August 2011 – Co-ordinating Editor: Manfred Finckh.

Introduction

Tropical forests are highly diverse ecosystems that play a key role in the global carbon cycle (Lewis 2006; Denman et al. 2007). Despite the importance of this biome there are few documented examples of databases which hold tropical forest data. Most of the vegetation databases registered in the Global Index of Vegetation-Plot Databases (GIVD; Dengler et al. 2011) hold information on relevés located in European countries.

Forestplots.net was developed primarily to hold information on pristine tropical forests; however woody savanna and logged forest plots are currently also included. The aims of the database are to a) enable network participants to access and manage their data online in a secure environment, b) to integrate and standardize forest plot data, c) to promote data sharing among the scientific community, and d) to provide access to publicly available forest inventory data. In this paper we review the key features of ForestPlots.net, and present the current limitations and future developments.

Scope

ForestPlots.net holds information on individual trees from plots in tropical and transitional zones monitored by more than 100 independent researchers, primarily members of the RAINFOR (http://www.rainfor.org; Mahli et al. 2002), AfriTRON (http://www.afritron.org; Lewis et al. 2009) and TROBIT (http://www.rainfor.org) networks. During plot establishment, all individual stems with ≥ 10 cm diameter are tagged and their diameter recorded (http://www.geog.leeds.ac.uk/projects/rainf or/pages/manuals_eng.html). Field campaigns to monitor diameter growth and stem turnover take place on average every three years. Currently Forestplots.net hosts more than 800 long-term plots. Approximately half of the plots in the database were established in the 1980s and 1990s. However, new plots are continuously being set-up and the database holds records for more than 250 plots established since 2000, mainly in the African continent. Presently 47% of the plots are located in Africa, followed by 43% in South America, 6% in Asia and 4% in Australasia. Plot size ranges from 800 to 50,000 m², with a mean plot size of 8,822 m². The percentage of plots by plot size category is as follows: 3% between 100 and 999 m²; 30% between 1,000 and 9,999 m², 61% between 10,000 and 19,000 m² and 6% ≥ 20,000 m².

Forestplots.net provided the data management environment for numerous papers developed by RAINFOR, AfriTRON, and TROBIT scientists, including several recent high-profile publications on the carbon balance and drought sensitivity of tropical forests (e.g., Lewis et al. 2009, Phillips et al. 2009, Phillips et al. 2010).

Technical specification

ForestPlots.net is a web application that can be accessed at http://www.forestplots.net. The front end was developed using the Microsoft .Net Framework, and it interacts with a Microsoft SQL server 2008 database hosted at Leeds University. Information about an earlier version of the database developed in Microsoft Access can be found in Peacock et al. (2007). The front end of the database is available in English, Spanish, Portuguese, and French.
Due to the complexity of the application and the potential for changing and increasing the functionality, a modular design was implemented. This design consists of five main components, which are the website front end; an application layer; a data control layer; a language controller, and a security controller. The application layer contains a presentation module and a set of functionality modules. An overview of how these components interact with each other and the underlying database is shown in Figure 1.

The website front end contains all the pages for the website and controls the way in which the user interacts with the site, and the application layer contains all the functionality of the web application. The presentation module contains a set of controls for formatting and displaying data to the user. The functionality modules allow new features to be added to the web application without the need for changes in other modules. Examples of functionality modules are the data import wizard and the query library. The data control layer provides and restricts access to the records in the database. It acts as a low level security barrier, and checks individual plot and census records against the credentials of the current user to permit or deny access to the record. Interactions with the database are handled using NHibernate, which maps the database tables onto a set of objects, helping to automate basic database operations and increase security by preventing website hacks such as SQL Injection. This is combined with automatic validation of objects before they are stored in the database, ensuring that data for all objects are within the expected range.

The language controller acts as a multilingual dictionary for the web application. Every piece of text within the system is identified by a keyword, and using this keyword, the corresponding phrase in the required language can be returned automatically. This allows new languages to be easily added by simply providing a new set of data files. The security controller provides information about the access rights of the current user. Both the language controller and security controller can be accessed from any layer within the system.

Database structure and data stored

The underlying database in ForestPlots.net is a relational database which utilizes more than 50 tables to store plot and individual tree information. The core tables of the database hold information on each individual plot (i.e. location, forest type and sampling protocol) and on the taxonomy and diameter measurements of each individual tree.

The table tblPlot is the key table holding plot information; the unique plot identifier (PlotID), Plot Code and Plot Name are saved in this table. Figure 2 shows all the items recorded in tblPlot and the tables that link to this table. The basic information that needs to be provided for each plot is: plot code, plot name, country, area (ha), as well as latitude and longitude. Additional information that can be stored in the database includes: forest type (i.e. based on edaphic characteristics, altitudinal location or composition), sampling protocol (minimum diameter size of measured trees if different to 10 cm), shape of the plot (i.e. square, long-transect, non-contiguous) and multiple GPS measurements of the perimeter of the plot or along transects.

Information at tree level is saved in table tblTree where each individual tree is assigned a unique identifier (Figure 3). This table is linked to tblPlot by the foreign key PlotID. The latest taxonomic information (SpeciesID), location in the plot (x, y) and TagID are saved in this ta-
Taxonomic information is saved in three tables, tblFamily, tblGenus and tblSpecies, which contain normalized information on family, genus and species. Family names follow the Angiosperm Phylogeny Group II flowering plants classification (APG2003) and the species list includes only accepted names from TROPICOS (http://www.tropicos.org), African Flowering Plants Database (http://www.ville-ge.ch/musinfo/bg/cjb/africa/recherche.php) and Queensland Herbarium (http://www.derm.qld.gov.au/wildlife-ecosystems/plants/queensland_herbarium/index.html). Taxonomic information will be revised against “The Plant List” (Kalwij in press; http://www.theplantlist.org) to update the database species master list. Information on the different tag numbers given to a particular tree is saved in tblTreeTag and the TagID of the current tag number is also stored in tblTree.

Census information for all plots is stored in tblPlotCensus and includes: census number, start and end dates of each field campaign and comments (Fig. 3). The table tblPlotCensus is linked to tblPersonnelPlotCensus, the later holds information on the people that participated in each field campaign, their level of involvement (personnel category; i.e. field leader, botanist, assistant) and the project that sponsored their participation. The most important table of the database is tblTreeCensus which holds information on the alive or dead status of each individual tree and the repeated measurements taken when alive during different field campaigns (Fig. 3). The table tblTreeCensus is linked to tblTree and tblPlotCensus with the foreign keys TreeID and PlotCensusID, respectively. For each individual tree sampled in the field, the database stores if it is alive or dead, its status, diameter, the height along the stem at which the diameter was measured (POM), measurement technique and data manipulation. Currently, tblTreeCensus holds more than a million records on individual measurements.

Plot data management

Each plot has a plot database manager which is the person who organizes census data uploads and oversees data edits and quality control. The plot data manager can make the plot metadata or one or more censuses publicly available. Access to plot information is granted on a plot-by-plot basis, however, for some plots, data from recent censuses is available publicly. In most cases only the Principal Investigator and associated members of his/her research group have access to a particular plot(s) which are shown when clicking in the “My Plots” section (Fig. 4). Additionally, all registered users have access to publicly available data found in the “Public Plots” section (Fig. 4), which they can view, download and use in accord with the ForestPlots.net user agreements.

Functionality

ForestPlots.net allows users, depending on the level of access, to view, edit, delete, upload, download and query data. View, edit and delete permissions can be granted at plot and plot-census level. Uploading, downloading and querying are performed at plot-census level, including information of all the trees in a plot at given census. Plot and census information of new plots can be uploaded with the “Plot Import Wizard”. Correspondingly, new census information of plots already in the database can be uploaded with the “Census Import Wizard”. Both “Plot Import” and “Census Import” wizards allow the user to upload one census at the time, coded following the detailed RAINFOR protocol for recording the state of all trees sampled in that census, whether alive or dead (http://www.rainfor.org/publications). The taxonomic information and the coding applied to each individual tree are verified and the user has the option of editing non-valid data at this stage. The upload wizards were developed to ensure a consistent methodology is followed whenever a plot is imported and to increase the quality and accuracy of the data.
Fig. 2: Diagram of the relations between the plot table (tblPlot) and associated tables. Primary keys are underlined. Items in bold are required. Relationship types: many optional \( \geq 0 \), required one \( 1 \), optional one \( o \).
Fig. 3: Diagram of the relations between the tables: tree (tblTree), plot census data (tblPlotCensus), tree census information (tblTreeCensus) and associated tables. Primary keys are underlined. Items in bold are required. Relationship types: many optional ➡️, required one ➡️, optional one ➡️.
When the uploading process is completed the application prompts the user to check plot stem count consistency and extreme diameter growth rates (> 40 mm yr\(^{-1}\) and 5 mm absolute decrease between two consecutive censuses) of individual trees by running the “Error Check Wizard”. At this stage the user can also run the “Automated POM Change” wizard to standardize diameter measurements of trees that had a change in point of measurement (“POM”) due to buttresses or other deformities. Post-field management of erroneous data and trees with POM change can impact basal area and biomass calculations at plot level, so we developed a methodology that can be applied consistently with different data sets. Further information on the “Error Check Wizard” and POM change can be found in Lopez-Gonzalez et al. (2011).

Plot data can be downloaded in two formats: “Plot Dump” and “Field Sheet”. Plot Dump files include all trees and their measurements and status in all censuses. Downloading this type of file is useful for analysing stem dynamics and calculating growth rates of a particular plot. The “Field Sheet” files, as their name indicates, are files formatted to be printed and used in the field. This file only includes the last three censuses (when available) and the headers for the information that needs to be recorded during the field campaign. After the field campaign, data can be entered into this file and uploaded as a new census. The “Field Sheet” files were developed to standardize the data gathering process during the field campaigns and to speed up data entry and uploading.

A novel feature of ForestPlots.net is the query library, which allows users to produce downloadable outputs on different query categories: biomass, basal area, stem dynamics, taxonomic information and wood productivity (Table 1). The database has two query libraries, one for plots found in “My Plots” and another for “Public Plots”. Both query libraries function in the same way, however the “Public Query Library” does not include queries that calculate change. Users can select the type of query they want to run from a drop-down box. Once the query type is selected, the plot(s) to be used need to be selected. The query parses the plot(s) identity before running and producing an output which can be viewed and/or downloaded. Full descriptions of the queries are displayed when users select a query to run. The query output downloadable file also includes the query description. Most of the queries use individual tree taxonomic information and diameter measurements as input and generate outputs on a plot-census basis. Table 1 summarizes the main query categories and their data inputs.

Due to the relevance of biomass calculation for the analysis of carbon stock dynamics we have included in the query library different types of allometric equations for calculation which are reported in Chambers et al. (2001), Baker et al. (2004) and Chave et al. (2005). Wood density is one of the parameters commonly used in the calculation of biomass; therefore we have incorporated the global wood density database (Chave et al. 2009, Zanne et al. 2009) into our database to assign wood density values to each individual tree.

**Limitations**

The main limitation of ForestPlots.net is the taxonomic data management. The application does not yet deal explicitly with synonyms, varieties, and morphospecies. ForestPlots.net is available online worldwide, but access can be difficult in remote tropical areas with poor internet access. While viewing and downloading of data are relatively easy tasks, users responsible for data uploading and error checking may need training sessions before undertaking these tasks.
Plate: Vegetation types featured by the vegetation-plot database GIVD 00-00-001.

A: Mature flood plains from Madre de Dios, North Peru (Photo: R. Brie- nen).

B: Amazonian moist forest, Acre Brazil (Photo: T. Feldpausch).

C: Savanna forest in Cameroon (Photo: T. Feldpausch).
**Future developments**

We are continuously working on improving ForestPlots.net. Our next aim is to expand the query library to include new biomass equations for tropical regions. We also plan to include some functional traits (i.e. height, leaf size) to complement the information already available in the database. We expect to include more plots in the database and to make an increasing fraction of the plot data publicly available. Other innovations are planned, including permitting, to improve ease of use and data visualization to support our aim of wider and deeper involvement by scientific and public communities in the tropical forest countries themselves.

**Acknowledgements**

Forestplots.net development has been supported primarily by the Gordon and Betty Moore Foundation in their grant to RAINFOR to understand the Amazon carbon balance. Additional support has been given by the School of Geography at the University of Leeds, as well as by grants from the Royal Society (S.L.L.) and from NERC (O.L.P.). We thank Julie Peacock, Timothy Baker, Ted Feldpausch, Roel Brienen and Lindsay Banin for their comments and suggestions.

**References**


