Facilitation of global change research in the tropics: science and data management

Alana Belcon, Liza Comita, Forest Isbell, Reynaldo Linares-Palomino, Carlos Rojas & Adam Wilson

OTS-PASI, May 2009

Introduction

To date, much global change research has focused on temperate and polar systems (Parmesan, 2006). However, it is becoming increasingly clear that even small changes in temperature and precipitation can have large consequences for organisms in the tropics (e.g. Clark, 2007). Since the mid-1970s, tropical regions have experienced a mean temperature increase of 0.26°C±0.05°C per decade (Malhi & Wright, 2004). Global climate models predict overall increases in tropical temperatures of 1–4°C in the 21st century and regional changes in precipitation with local increases of up to 8% and local decreases of up to 35% (IPCC, 2007). While it is widely acknowledged that these shifts in climatic conditions will likely have significant effects on the flora and fauna of the tropics, there is currently insufficient data to predict and mitigate these impacts. In addition to climate change, tropical ecosystems also face a host of threats including deforestation, pollutants, habitat loss, overexploitation, degradation and fragmentation (Wright, 2005). Given that tropical ecosystems provide a wealth of ecosystem services and harbor a large proportion of the earth's biodiversity, urgent action is needed to increase understanding of how these valuable and complex systems will respond to ongoing and predicted threats.

Studies of global change in the tropics have been hindered by several challenges unique to tropical systems. First, the high species richness of most biological communities in the tropics makes it difficult to carry out inventories and to characterize interactions. As a result, we lack basic data for most tropical species, including species' resource requirements, environmental tolerances and distribution patterns. Thus, it is not surprising that we lack sufficient information to make informed predictions of how tropical species and communities will respond to changes in climatic conditions or the degree to which they can recover from other types of human disturbance. Second, most countries in the tropics are developing nations and lack the infrastructure and funding for large scale global change research. As a result, there is currently a paucity of long term climate and biological records for most tropical sites, making it difficult to ascertain whether communities are changing and whether such changes are due to changing environmental conditions.

With these challenges in mind, in this paper we explore the types of data needed to address effects of global change on tropical systems, the nature of funding required for tropical global change studies, the protocols and methods for collecting needed data, and the storage and dissemination of data to allow access, both now and in the future, to researchers working on global change in the tropics (Fig. 1).

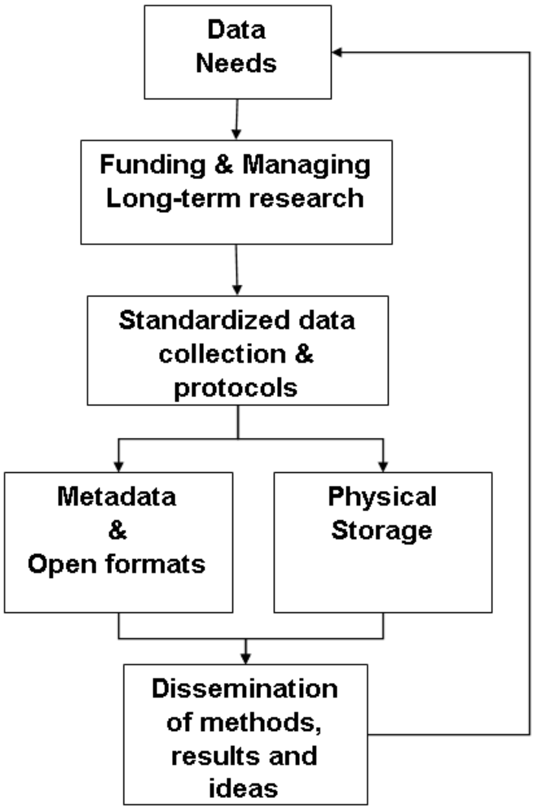

Figure 1. Flow chart of science and data management. Each box represents a stage in the management of science and data to study global change in the tropics.

Data for addressing global change

There are many unanswered questions concerning how tropical biota will respond to global change. For example, we currently have a poor understanding of the role of tropical forests and oceans in carbon sequestration, the ability of tropical species to adapt to rapidly changing environmental conditions, the number of species contained in secondary versus old growth forests, or how species assemblages and interactions will change under novel climatic conditions.

Published studies on effects of global change in tropical ecosystems have largely been the result of individual researchers noticing shifts at their study site after the fact, rather than from pre-planned studies specifically designed to examine effects of climatic changes. For example, there have been observations of shifts in species distributions in Monteverde (Pounds et al. 1999), declines in reptile and amphibian abundances (Whitfield et al. 2007) and changes in tropical tree growth rates at La Selva (Clark et al. 2003). While these studies have documented important changes, such observational studies often result in little consensus about general patterns due to differences in site histories or research methodologies. For example, there is ongoing debate over whether tropical tree demographic rates are increasing or decreasing in response increase atmospheric C02 (Clark 2004).

Furthermore, such studies provide little information concerning the mechanisms behind observed changes. For instance, there is debate over the relative impacts of climate change, chytrid fungus, and other factors in causing amphibian declines (Pounds et al. 2006, Young et al. 2001).

We proposed that these problems can be remedied by several types of studies: (1) long-term monitoring using standardized protocols, (2) observational studies of species biology and interactions under current conditions and (3) experimental manipulations to assess species' responses to environmental variables.

Long-term monitoring

Long term climate records for research sites are clearly needed in order to address issues of climate change (Lovett et. al. 2007). Collecting these data should be a top priority. Measurements of minimum and maximum temperature, total precipitation, evaporation, relative humidity, and soil moisture should be collected on a daily basis, or more frequently, at tropical research sites worldwide. Few such data sets currently exist for research stations in the tropics. In the future, the availability of automated weather stations should make the collection of these data much easier (see section on methods and protocols). At sites lacking historical records from long-term monitoring, climate proxies could potentially be used to reconstruct rainfall or temperature records. These proxies include tree growth rings, coral cores, stalagtites and stalagmites, and river and lake sediments.

In addition to monitoring climate variables, there is also a need for long-term monitoring of land use patterns in order to document changes and to understand the larger spatial matrix in which natural communities are situated. Recent advances in remote sensing offer researchers the opportunity to track land use changes accurately and easily (Clark et al. 2004).

Long-term data on tropical flora and fauna are also needed in order to document impacts of global change. Specifically, studies are needed that monitor species abundances, dynamics, and distributions. The high diversity of tropical ecosystems precludes monitoring all species. Therefore, researchers will have to be selective, perhaps relying on indicator species across multiple sites, or concentrating efforts at sites that are particularly at risk or are biodiversity hotspots. Monitoring schemes should consider which spatial and temporal scales are most relevant to the organism(s) under study and to the threats to those organisms (Ferraz et al. 2009). For example, most inventories of tropical forests have been done at 2-5 year intervals, making it difficult to determine how annual changes in temperature or precipitation affect dynamics (Clark 2007).

On a practical level, long term monitoring requires continuity. Projects should involve both senior and junior PIs, with leadership transitioning smoothly from one generation to the next. In addition, there is also a need for dedicated, trained project personnel, and care should be taken to insure that new technicians are trained properly. As we discuss further below, long-term monitoring efforts are often hampered by a lack of secure and sufficient funding. In the absence of large grants, researchers may turn to citizen scientists or work with school teachers to have students help collect data. Researchers should also consider training non-academic, local community members to aid in monitoring efforts (Ferraz et al. 2009).

Short term studies

While long term monitoring is essential for documenting impacts of global change, more immediate studies can be undertaken that will enhance our ability to predict when and where impacts will be greatest. This will involve both observational and experimental studies:

Observational studies

Due to the high diversity of tropical communities, detailed information is available for few species. In addition, at less well-studied sites, there may be a number of species that have yet to be described. Thus, there is a need to catalogue and study tropical organisms, as well as interactions between organisms, under current conditions, preferably before they are more severely impacted or lost due to anthropogenic impacts. In other words, for most tropical systems, we need to establish a baseline against which future changes can be compared. However, researcher must keep in mind that sites may have already been impact by humans. Clark (2007) points out the need to document past disturbance at research sites in order to understand how current changes are affecting tropical species, and recommends studies of site history for the past 500-1000 years. Such studies involve sampling for phytoliths and charcoal, and analyzing the flora for patterns that suggest previous human or natural disturbance (Clark 2007).

Other short-term observational studies that would help advance global change research are studies of species distributions and dynamics along temperature and precipitation gradients. These studies will help identify what limits species ranges, and may serve as a proxy for species' environmental tolerances in the absence of experimental manipulations (see below). These sorts of observational studies may turn in to long-term studies, but should be designed to yield pertinent information in the short-term

Experimental studies

There have been few attempts to address tropical species' sensitivity to environmental variables relevant to global change. For example, "the heat sensitivity of tropical lowland plant species is an open question" (Svenning & Condit 2008). This lack of basic information hampers efforts to forecast impacts of global change on biodiversity in the tropics (Svenning & Condit 2008). Thus, there is a pressing need for studies of physiological mechanisms underlying climate effects on species and ecological communities (Portner et al. 2009).

For organisms that can easily be grown or reared in the lab, such studies would involve using controlled environments (i.e. growth chambers, greenhouses, microcosms) to tease apart effects of environmental variables on species' performance. Of particular interest are species' reactions to temperature, light, C02 and water availability. Field studies, involving transplants or manipulation of environmental variables (e.g. irrigation experiments, C02 addition) will also provide valuable information about how species' performance and interactions vary with climatic conditions.

Species' vulnerability to climate change will also depend on whether it possesses mechanisms for either ecological or evolutionary adaptation to changing conditions (Williams et al. 2008). Currently, there is no information on tropical species' ability to acclimate to changing climate and whether sufficient genetic diversity exists within species' to allow for adaptation to novel climates. In addition, studies of species' dispersal abilities are needed to determine whether species can migrate to higher elevations or latitudes at rates fast enough to avoid extinction.

Computer simulations can also serve as another type of experimental study in which species and community responses to changing conditions can be assessed. Such studies would involve modeling changes in species distributions, community composition or diversity using climate change predictions compared with current species distribution (e.g. Colwell et al. 2008), and preferably would incorporate a mechanistic understanding of the determinants of species range limits (Svenning & Condit 2008).

Funding & managing long-term research

After identifying data needs, how will we fund global change research in tropical ecosystems? Currently, it seems that most of the funds supporting this research come from the US National Science Foundation (NSF). All NSF grants require that at least one of the principle investigators on a project are affiliated with a US institution, and most grants last only three to five years (NSF, 2009). Several European countries also allocate a relatively large proportion of their gross domestic product to research (King, 2004). Consequently, thus far, much of the global change research in tropical ecosystems has consisted of relatively short-term studies conducted by investigators who come from developed nations in temperate ecosystems. This may not be ideal.

We conducted a survey to determine how to improve long-term global change research in tropical ecosystems. All of the participants in our NSF Pan-American Advanced Studies Institute (PASI) course completed the ten question survey (Appendix 1). The first five questions were related to funding; questions six through nine were related to collecting, storing, and sharing data; and question ten allowed participants to offer any additional comments. Similar questions were asked in personal interviews with Zak Zahawi, station director at the Las Cruces Organization for Tropical Studies (OTS) site, and the PASI course coordinators: Deedra McClearn, Delphine Farmer, and Paul Foster. The survey results reported here only include the responses from participants that were not in the research and data management PASI working group.

Interestingly, the survey results indicated a general consensus that long-term global change research in tropical ecosystems should become increasingly globalized. Most participants would like to see this research be funded by a global pool of money from many governments and nongovernmental organizations (Fig. 2a). Most participants also indicated that collaboration between investigators in temperate, developed countries and tropical, developing countries is ideal and should be required (Fig. 2c). Additionally, most participants would like to see longer funding cycles (Fig. 2d, e), more standardized protocols (Fig. 2f), and a global data storage system (Fig. 2g).

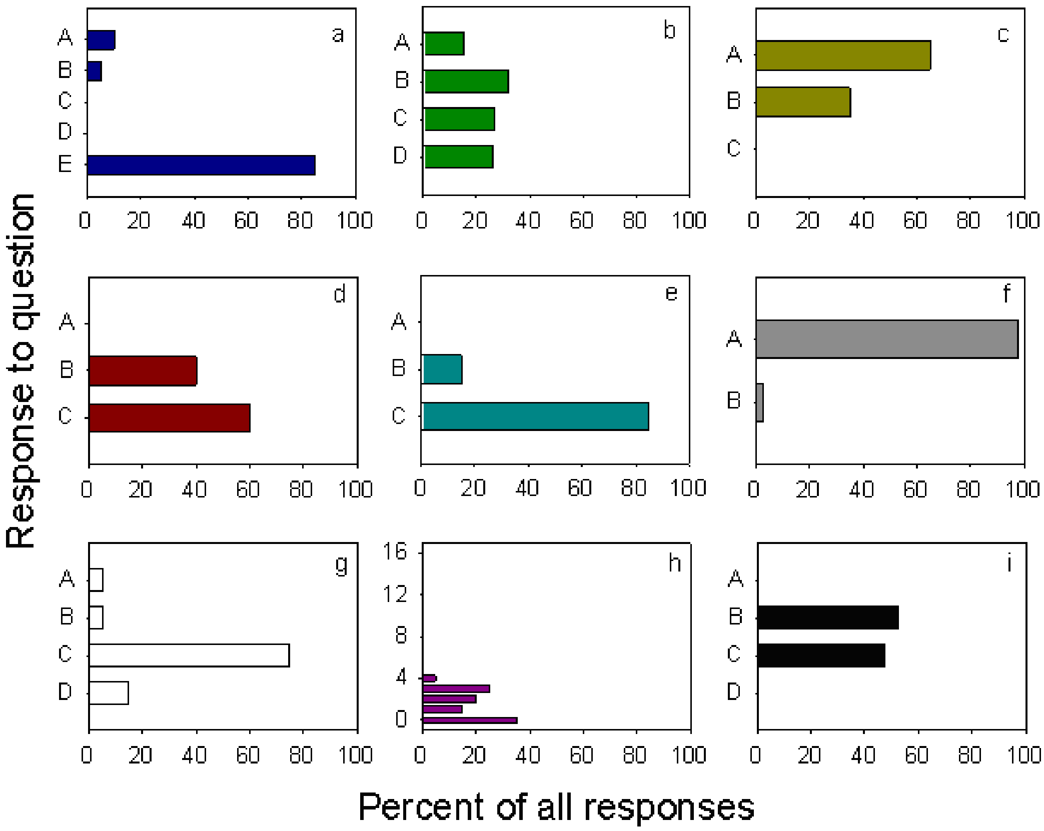

Figure 2. Responses of course participants to the first nine survey questions. Panels a-i correspond to questions 1-9, respectively, on the survey (see Appendix 1 for questions). In panel h, the number on the y-axis is the number of collections (out of 16) that students were familiar with.

The interviews also suggested that long-term global change research in tropical ecosystems should become increasingly globalized. Interviewees stated that that the majority of the long-term funding currently comes from NSF, and that most of the investigators come from the US and Europe. They also stated that there should be increased opportunities for collaboration between investigators in tropical developing countries, and temperate, developed countries. This suggestion is supported by the findings that multi-authored papers have a higher citation rate than single-authored papers, especially when there is interdisciplinary collaboration between institutions (Leimu & Koricheva, 2005).

Globalization of research in this field will likely accelerate our understanding of the causes and consequences of global changes in tropical ecosystems. At present, there are relatively few funding opportunities for investigators in the tropics (King, 2004), and few funding programs that are structured according to the ideal that was indicated in the survey and interview responses. However, there are examples of successful funding programs that are similar to this ideal. For example, the Natural Environmental Research Council in the United Kingdom has an Ecosystem Services for Poverty Alleviation program that required partnerships between investigators in developed and developing nations (NERC, 2008). This brings together the financial resources and infrastructure available to the investigators in the developed nation with the long-term local experiences and personal interactions of the investigators in the developing nation. We suggest that a similar funding program could be developed for long-term global change research in tropical ecosystems.

Standardized data collection and protocols

Tropical forests store ca. 40% of the carbon residing in terrestrial vegetation and harbour more than 50% of the world's species. Relatively small yet consistent changes within the remaining tropical forests could have global consequences for the climate, biodiversity, the global C cycle, the rate of climate change and hence human welfare (Lewis *et al.* 2004). Thus, long-term monitoring of plant and animal populations and functions are of relevance if we are to understand and disentangle natural, temporal changes (e.g. ontogenetic growth changes of individual tree species) from changes that may be caused by external factors such as changes in land use and climate (e.g. growth variation of tropical forests as a response to climatic/atmospheric variations) (Chave *et al.* 2008; Clark 2007).

There are few undisturbed ecosystems left on earth and long-term monitoring programs in these areas are of increasing value, not only because they may provide information about the functioning of ecosystems, but also because they can be used as a baseline when comparing with other disturbed areas (Spellerberg, 2005). Several working groups and networks exist which aim to explore and analyze long-term effects of global change on tropical forest ecosystems (Fig. 3).

Most of these sites are using detailed standardized protocols for monitoring factors and variables of interest to these working groups, usually developed after several years of adjusting and perfecting (Phillips et al. 2008). However, to ensure exchange of information, as well as to compare the outcomes of long-term research and its inclusion in meta-analysis across different ecosystems and continents, there is a need for methodological uniformization across users and projects., as well as communication and collaboration among workers in different disciplines and at different sites (Condit, 1995; Druce *et al.* 2008). In addition to this issue, some other relevant aspects emerge such as the spatial and temporal extent of the survey. Indeed, there has been an ongoing discussion about the contrasting results obtained from a few but globally distributed large-scale forest inventory plots (16-50 ha in size) influenced by different climate and other physical and social variables, compared to wide-spread smaller plots located mainly in Amazonia. For instance, Baker *et al.* (2004) inferred, a long-term increase in aboveground biomass from the Amazonian dataset, whereas Chave *et al.* (2003) and Laurance *et al.* (2004), found no change in aboveground biomass/basal area when they analyzed part of the large plot dataset. In any case, as has been advocated by Chave *et al.* (2008), these efforts should be viewed as complementary rather than competing, especially since they constitute some of the few networks in the tropics that have already a couple of decades worth of data. Another issue that needs to be addressed is that of the frequency and timing of data recording. D. Clark and D. Clark (RCN workshop presentation, see also Clark, 2007) have found that most long-term research to study climatic responses across tropical forests has census intervals of 2.6 years (62 study sites) and nearly all current monitoring is based on multi-year intervals. This is not enough to record changes adequately and they propose to record annual measurements.

New and promising data acquisition methods

This section will focus mostly on recent developments in remote sensing technologies that are allowing the analysis of several tropical forest characteristics over greater spatial scales and over longer time periods but with higher accuracy. Key advances that are allowing this include the development of new and more powerful sensors, image-processing techniques and the comparison and integration with field data (Chambers *et al.* 2007). These new technologies, algorithms and collaborations between ecologists and remote sensing specialists are allowing the study of ecosystem processes and forest function (e.g. phenological cycles, canopy chemistry), forest structure and species composition (e.g. crown detection and species identification, estimation of forest biomass and detection of changes in forest structure over large heterogeneous areas) and anthropogenic impacts (land-use and land-cover changes, like better detection of diffues forest disturbances like selective logging practices) (Chambers *et al.* 2007) (Fig. 4).

RAINFOR (Amazon Forest Inventory Network, http://www.geog.leeds.ac.uk/projects/rainfor/) is an international network that has been established to understand the biomass and dynamics of Amazonian forests. RAINFOR was established to bring together researchers throughout Amazonia who maintain small permanent forest sample plots (usually of 0.1ha and 1ha). By compiling and comparing these studies on a regional scale they were able to obtain information that provided insights into the mechanisms underlying the current responses of Amazonian ecosystems to climate and the possible future of Amazonia under global change scenarios.

The Center for Tropical Forest Science (CTFS, http://www.ctfs.si.edu/) is a global network of over 30 forest research plots (located in the Americas, Africa, Asia, and Europe, with a strong focus on tropical regions) committed to the study of tropical and temperate forest function and diversity. CTFS conducts long-term large-scale research on forests around the world to increase scientific understanding of forest ecosystems, to inform sustainable forest management and natural-resource policy and to build capacity in forest science. Common plot structure and scientific methodology unify the CTFS network. Because each plot follows the same methodology, scientists can directly compare data collected from different forests around the world and detect patterns that would otherwise be impossible to recognize.

TEAM (Tropical Ecology, Assessment, and Monitoring, http://www.teamnetwork.org/en/home) is a network set up by Conservation International to monitor long-term trends in biodiversity through a network of tropical field stations with the aim to provide an early warning system on the status of biodiversity that can effectively guide conservation action at local, regional and global scales. The TEAM initiative is in many ways different to the former two networks, because it focuses not only on vegetation, but has a more integral monitoring program including additionally birds, butterflies, terrestrial vertebrates (camera trapping), primates, amphibians and climate. They do this by monitoring several indicators, including biomass, abundance and diversity of species, habitat size and health, and the rate and delivery of ecosystem services to people. An important aspect of their work is the generation of information to understand the dynamics of biodiversity and its responses to human-induced changes such as habitat conversion, habitat degradation, overexploitation of species, and climate change.

There are several other tropical long-term initiatives like the Carbono Project at La Selva (http://www.fiu.edu/~carbono/tower.htm) or some of the research sites of the International Long-term Ecological Research network (http://www.ilternet.edu/), such as those in Mexico,Venezuela, Brazil, Thailand and Australia.

Figure 3. Some examples of long-term research networks in the tropics using standardized methodologies.

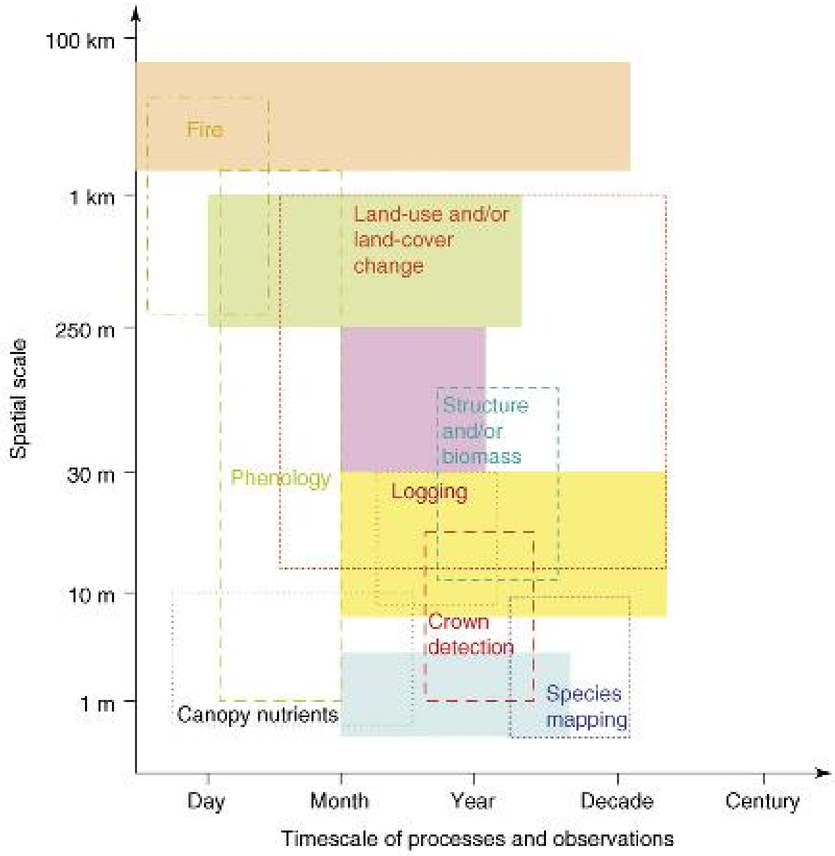

Figure 4. Scales and sensors that are improving our understanding of tropical forest ecology (Chambers *et al.* 2007).

Metadata & Open Standards

Global change research requires long term datasets to understand and contextualize observed ecosystem changes. These include meteorological data as well as ecological studies that are often not designed for long term research. These data are usually plagued by inconsistent methods, changing storage medium and hardware corruption, missing data, and other issues mentioned in previous sections. This leads to a loss of the information content of the data over time (Fig. 5 and Michener *et al.* 2007). This loss of information diminishes our ability to detect long-term trends in ecosystems and so effort should be made at the time of data collection to record and store the associated metadata. Fortunately, there is a growing trend in making data and methods publicly available shortly (2-5 years) after data collection (Clark, 2007). The challenge is in organizing this growing quantity of data to maximize its utility in the future.

Metadata

Detailed metadata preserve the information needed for researchers to use data in the future to continue studies and use the data for novel purposes. Ecological Metadata Language (EML) is an example of a flexible, structured format in which to store metadata and has been endorsed by several large research organizations, but is not used in many tropical areas

(Table 1). Developing metadata can be time consuming but is important to provide the information necessary to preserve the utility of the data for a longer period of time, to continue previous studies, and to put the data to new uses. Complete metadata also make the data useful to future researchers that have no direct knowledge of the context in which the data were collected.

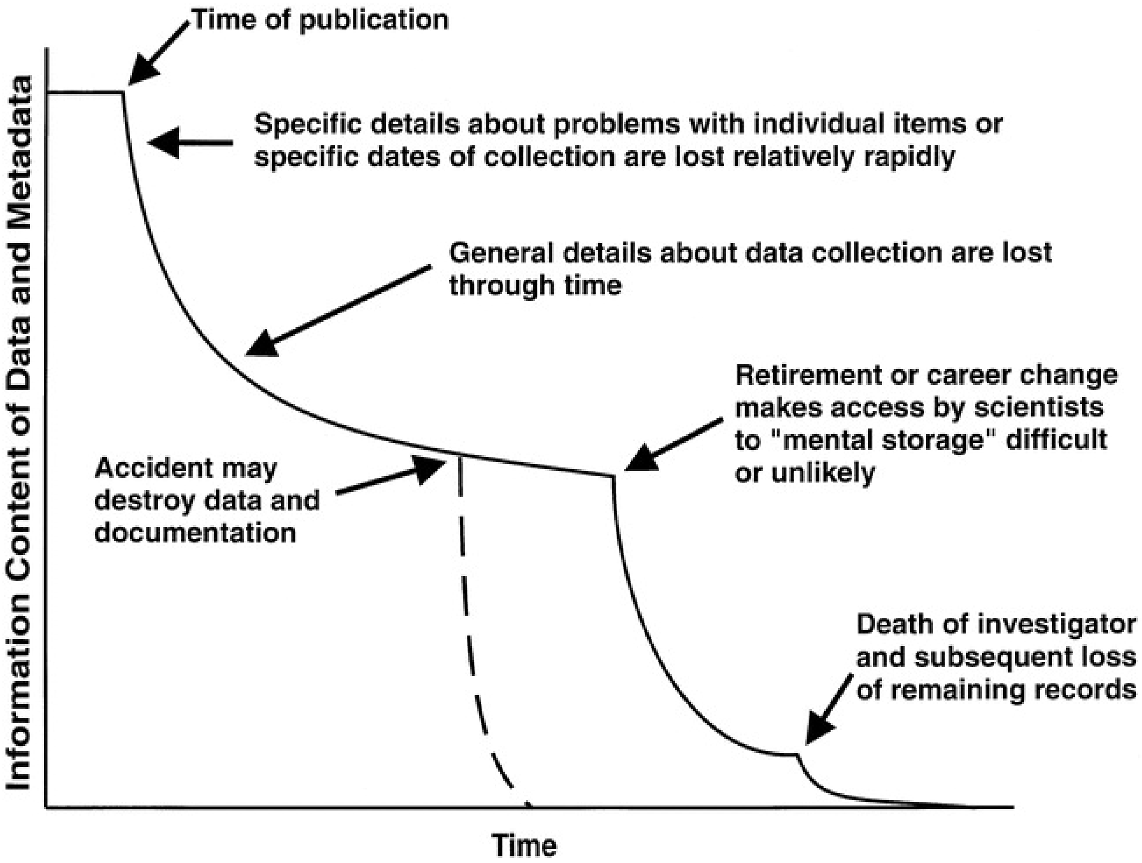

Figure 5. Example of the normal degradation in information content associated with data and metadata over time ("information entropy"). Accidents or changes in storage technology (dashed line) may eliminate access to remaining raw data and metadata at any time. Graphic from Michener, 1997.

The Knowledge Network for Biocomplexity (KNB) has made software available to assist scientists in developing metadata files in EML (Morpho Software, http://knb.ecoinformatics.org/morphoportal.jsp, accessed on May 7, 2009). The KNB also offers a free online storage of metadata (using the METACAT software) for ecologists (http://knb.ecoinformatics.org/index.jsp, accessed on May 7, 2009). Researchers should strive to archive complete metadata in publicly available databases such as those listed in the table below.

Data formatting

In addition to a standardized metadata format, it is also important to use open and 'industry-standard' data formats (such as ASCII files and other open formats) because they are less likely to become obsolete than binary proprietary formats such as Microsoft Excel (Conley & Brunt, 1991). In addition, long term storage of data in standard data formats facilitates access by those without specific (and often proprietary) software.

Table 1: Organizations that use Ecological Metadata Language to store and organize their metadata (KNB in the Community, http://knb.ecoinformatics.org/community.jsp, accessed May 7, 2009). Note that organizations are concentrated in temperate regions.

Knowledge Network for Biocomplexity (KNB)
National Center for Ecological Analysis and Synthesis (NCEAS)
U.S. Long Term Ecological Research Network (LTER)
International Long Term Ecological Research Network (ILTER)
Partnership for Interdisciplinary Studies of Coastal Oceans (PISCO)
Organization of Biological Field Stations (OBFS)
University of California Natural Reserve System (UC NRS)
Ecological Society of America (ESA)
Kruger National Park (KNP), South Africa
Long-term Individual based Time Series (LITS) project, UK
Electronic Tagging Data Repository, Pelagic Fisheries Research Program (PFRP)
Windsor-Essex Environmental Metadata System (WEEMS), Ontario, Canada
Taiwan Ecological Research Network (TERN)
Taiwan Forestry Research Institute (TFRI)
Borer Seabloom Lab Data Repository, Oregon State University (OSU)
Japanese Long Term Ecological Research Network (JaLTER)

Preservation of analytical processes

In addition to the data and metadata, it is also useful for researchers to preserve their analytical procedures for future use and/or comparison with current methods. This is often more complicated than preserving the data, but can be achieved by archiving the detailed steps necessary to process the data. One example of an effort to standardize analytical procedures is the Kepler analytical environment, which can organize data, run analysis using various tools (such as R, Matlab, and others), and generate output (http://kepler-project.org/, accessed May 7, 2009). This allows researches to easily share data and analytical procedures with each other and archive processes as well as products.

More sincere effort needs to taken on the part of researchers to compile complete metadata, data, and analytical processes. This combination of procedures results in a ‘package’ of data collection protocol, data with metadata, analytical software, and scripts that anyone (with access to a computer) can run and ensures that your analysis will be repeatable and useful into the future. It also allows the process to be continued through personnel changes or periods of no data collection.

Physical Storage and Data Accessibility

Even as we address the need for increased scientific knowledge about the interactions and processes that control tropical ecosystems we also struggle with storing and accessing that data. Increasingly large amounts of data are being produced each year, more than we are able to store. In his 2008 white paper, ‘The Expanding Digital Universe’, John Gantz, Chief Research Officer and Senior Vice President of IDC Corporation describes how the amount of information produced by the planet surpassed our storage capacity in 2007 and calculates that by 2011 we’ll be producing 10 times the amount of data we currently produce. While not all information that's created is stored, for the first time we are in the position where we could not store everything we created even if we wanted to and it will become increasingly neccessary to balance data storage redundancy with available storage space.

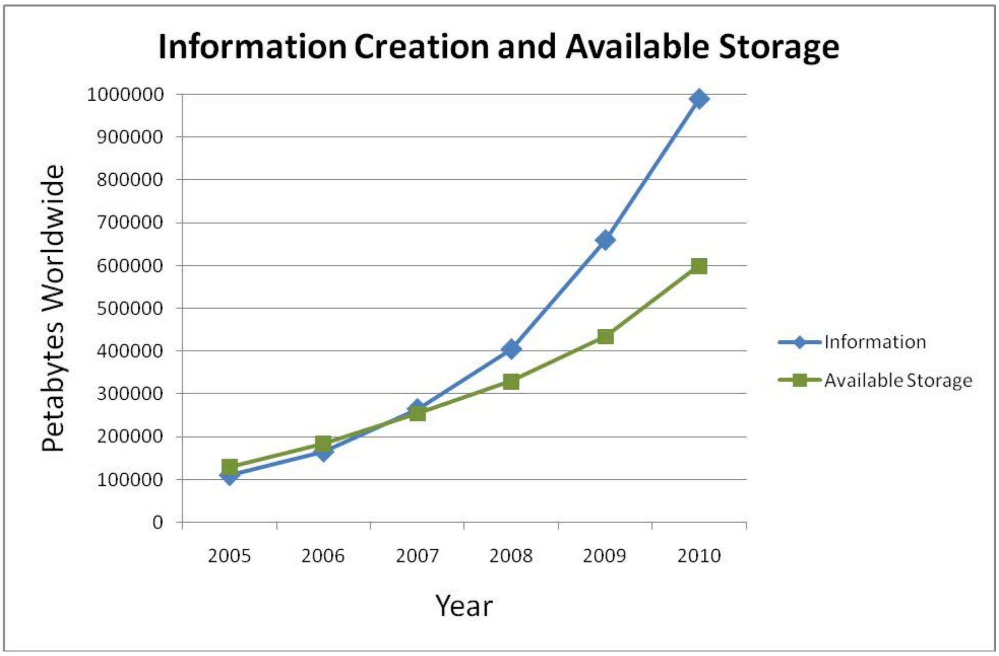

Figure 6. Information creation and available storage. Source: John Gantz, IDC Corporation: The Expanding Digital Universe

In addition, technological progress is occurring so rapidly that data file formats are becoming obsolete faster than data can be transferred to new formats, meaning that we are losing old, possibly valuable data. And perhaps most importantly, scientific data and the analyzed products are often stored in non-connected archives making access difficult especially across disciplinaries. Jones et al in their 2006 paper, The New Bioinformatics, discuss the need for ecologists to move away from the traditional spreadsheet–based tool or even the more robust relational database systems such as Microsoft Access. These databases are typically structured for the particular scientist’s needs and do not follow a common format making it difficult for other scientists to use as well as being difficult to access. One solution to the problem of project-specific databases complicating data integration is to develop vertically integrated databases that store data collected by many different investigators, but all following a common theme. Using these databases will i) demand a common format making application to other projects more viable and ii) provide easy access to other scientists. However the success of this method is dependent on each database being common knowledge and on funding/support to maintain it. Jones lists 16 metadata and data collections that are frequently used in ecological research that we presented to the OTS/PASI 2009 participants and requested that they indicate which collections they knew about. The results are shown below (Fig. 7).

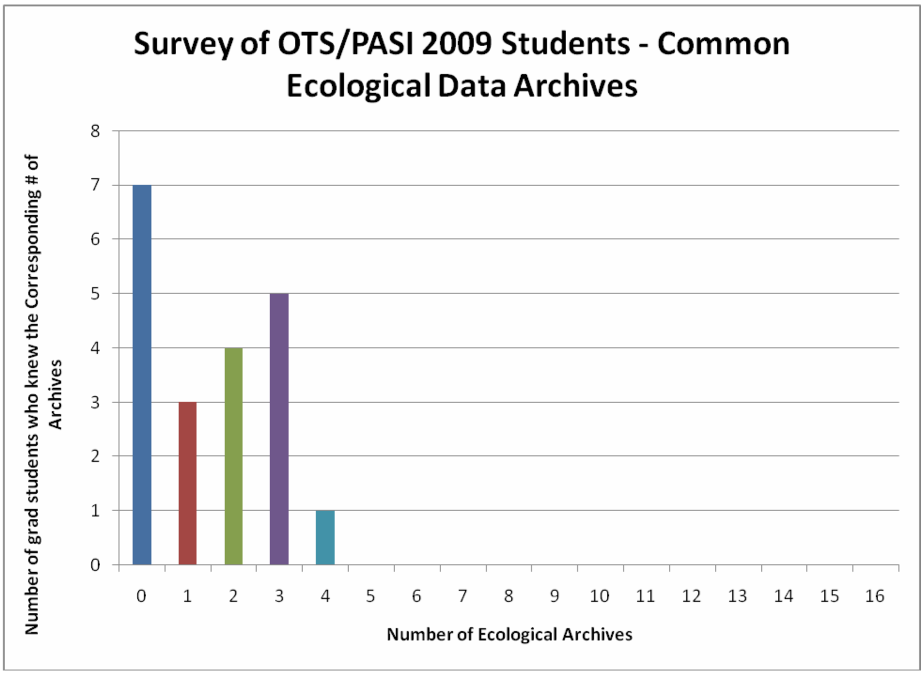

Figure 7. Common ecological data archives known by the OTS/PASI 2009 students

Twenty participants answered the survey and of those 7 or 35% did not know a single one of the ecological archives listed. Of the remaining 13 participants most knew only 1-3 archives and only one knew 4 of the 16 archives. Clearly simply having metadata and data archives within each subject is not enough if researchers do not know about them and therefore are not using them. One possible resolution for this is the creation of a mega archive that would link all others across disciplinaries. Bill Michener of the University of New Mexico is currently proposing the formation of 'Data One', a new cyberinfrastructure made up of nodes around the world that would integrate existing data from biological archives, individual scientists, citizen scientists, global and regional land cover change data etc into one gigantic network. Searching for 'Broad-billed Motmot' for example would result in links about the bird from scientific journals, citizen observations, GIS maps etc. Users would not have to access multiple archives and data accessibility should increase. Of course there would be the same initial problem of getting the word out about the network and encouraging people to actually use it. Some have suggested that one way to get around the creation of networks that individuals and institutions do not use is to have funding institutions form a loose network for data management and require that researchers submit their data within a specified period of time. The National Science Foundation currently requires this but does not truly enforce it. If they and other funding institutions were to enforce the submission of data that then became publicly available not only would data be easily collected into one network but scientists would know exactly where to go to search for and request data.

On the other hand, not all the news is bleak. Never before has scientific knowledge been as readily available as it is today especially to the average citizen. With the advances in modern technology interested individuals can download compilations of bird songs to their mp3 players and take with them into the forest to assist in bird identification while park rangers and forest guides can load entire digital herbariums onto palm-held devices and use in the field for rapid identification of plant species. In 2004 Daniel Janzen wrote in Philosophical Transactions of the Royal Society:

> *Imagine a world where every child's backpack, every farmer's pocket, every doctor's office and every biologist's belt has a gadget the size of a cell phone. For free. Pop off a leg, pluck a tuft of hair, pinch a piece of leaf, swat a mosquito, and stick it in on a tuft of toilet tissue. One minute later the screen says Periplaneta americana, Canis familiaris, Quercus virginiana, or West Nile Virus in Culex pipiens.*

Sounds like an impossible dream yet Apple Inc. in May 2009 demonstrated the prototype for their iPhone field guide where researchers in the field can now take pictures of leaves on their phone and have them digitally compared to thousands of species to provide a identification match. While a long way from Janzen's vision Apple's application demonstrates how rapidly technological advances are not only increasing our ability to disseminate data but is also placing scientific data into the hands of the layperson.

Data release, database regulation and limitations of the release process

Once the process of data storage is carried out, data release is the following step in the management of databases. This is an important aspect of the data cycle. Presently, some important funding agencies such as NSF and European counterparts (e.g. CSIC in Spain, DFG in Germany, CNRS in France) require PI's to make data available to the general public. These agencies consider data sharing a very important aspect of the scientific process due the social responsibility that the scientific community has with the surrounding society. However, the regulations concerning database protection are still under discussion at the international level. One example of this is the point of discrepancy between the United States and the European Union in relation to database protection (see DiGiusto, 2005).

Fortunately, there are other levels of the process in which there seems to be a consensus among some institutions. One of these cases is the publication of raw datasets supporting publications as a way to release the data faster. For example, instances in which results can be published quickly regardless of publication of actual scientific articles include *Ecological Archives* in the United States and the initiative called *Parse Insight* in Europe. This is an important issue; modern topics in science (e.g. climate change) require prompt availability of information. However, due a classical vision in which the performance of the scientific work is measured in terms of published articles, there is difficulty in doing this. Apparently, some ecologists even consider that one of the aspects that should be taken into account is that within this field data collecting is not seen as valuable as data analysis and production of scientific articles. Interestingly, a group of major stakeholders within the scientific community in the European Union are trying to lead the scientific world towards a complete open access system of data management. They are organized in the *Alliance for Permanent Access*, in which they developed a model for scientific information to be released in a dynamic way that represents the interest of the user (Fig. 8) and their logo of ***use it or lose it***. Similarly, the initiative of the *Biodiversity Commons* proposed by Thomas Moritz of the American Museum of Natural History (Moritz 2004a) pretends to promote the concept of the public domain and free access of biological information through this system of Internet-based regulations.

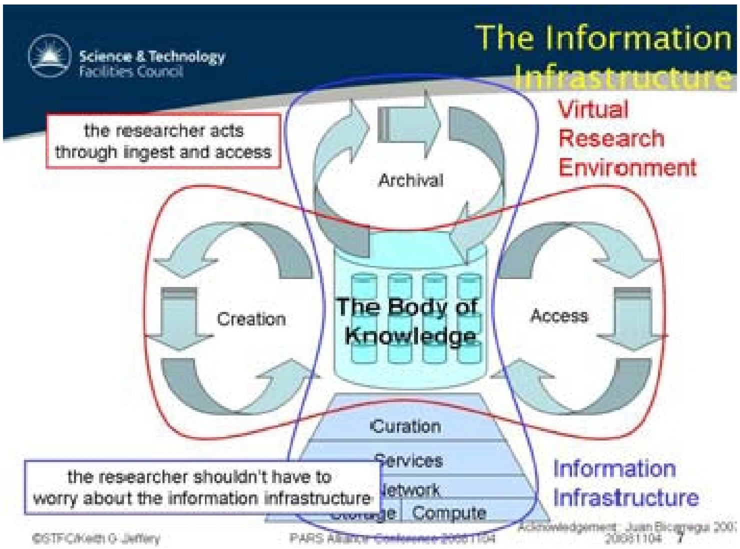

Figure 8. Diagram showing the data management proposed by the Alliance for Permanent Access in June 2008.

A remarkable aspect of these situations is the validity of storing collected data versus storing published information. This is very important for long-term research projects for which the possibility of publishing very similar analyzed data on a regular basis decreases with time. One of the problems associated with this, however, might come from a different angle. Once the collected data is released to the scientific community over the Internet, it seems that there are no clear international stipulations regarding the protection of authors. Resources on the Internet are still trying to be regulated by international legislation (Trembly, 2009). In the United States for example, databases are considered compilations of data and are therefore protected by the Copyright Act, a situation that occurs in many other countries. However institutions that make databases available to the scientific community have to deal with this obstacle and use different approaches. Once again there seems to be very few regulations in relation with this issue.

The Global Biodiversity Information Facility (GBIF, http://www.gbif.org, accessed May 9, 2009), hosted by the University of Copenhagen in Denmark is one example of that. They manage and store around 15 million records from more than 280 providers around the world. When data is given to them, the provider agrees on complying with certain regulations. One of them is the so-called Data Sharing Agreement that clearly expresses:

> *The participants who have signed the Memorandum of Understanding have expressed their willingness to make biodiversity data available through their*

nodes to foster scientific research development internationally and to support the public use of these data.

Interestingly, in the chapter on Intellectual Property Rights to Biodiversity Data the same document stipulates:

GBIF promotes the free dissemination of biodiversity data and, in particular should seek, to the greatest extent possible, to make freely and openly available, with the least possible restrictions on reuse, any data commissioned, created or developed directly by GBIF.

It is clear that the provider of information is the ultimate responsible for the use of stored data. In this way, GBIF evades any potential misuse of information on their servers. A similar approach is taken by the Inter American Biodiversity Information Network (IABIN, http://www.iabin.net, accessed May 10, 2009). In contrast, the Spatial Analysis of Local Vegetation Across Scales (SALVIAS, http://www.salvias.net, accessed May 9, 2009) project takes a different direction to protect data providers. Even though the structure of data management is similar to GBIF or IABIN in terms of participants involved in the process (e.g. providers and data manager), this project has a hierarchical structure in which permissions are granted for users to access different levels of information depending on their level. Their Intellectual Property Rights Policy establishes that:

The General Participant Agreement stipulates that data providers have automatic rights to co-authorship on publications resulting from analysis of their data. Anyone using data obtained through SALVIAS must adhere to both the General Participant Agreement, as well as the specific data access conditions set by the data provider.

Given all these conditions, it seems that even though there is no consensus on the topics, there are approaches that can be taken to protect data collectors. Unfortunately, all these efforts are useless if the information cannot be released.

In some developing countries, the scientific infrastructure is not developed enough to provide long-term support to databases and assure that the process of data availability takes place with normality. Institutional limitations play a role in this process as well. For example, some institutions are structured in a way in which the use of particular research data is considered institutional and not personal. This institutional use of data limits the researcher in relation to the type of information that can be released to the scientific community at a specific time during the research process. Even though in the modern world this is more common at the molecular biology research level, there are institutions in developing areas of the world in which this is the normal pattern for the use of information.

An important aspect that cannot be forgotten is the trend on the use of information that Internet resources have been undergoing during the last years. The Internet is no longer an enormous place to store information but a data platform in which users participate actively on the modification of these resources. For that reason, the 2008 Report of the Research Information Network (a consortium of scientific institutions in the United Kingdom) establishes as one of the conclusions that:

Researchers, funders, institutions, publishers and other stakeholders should *monitor the development and take-up by researchers of Web 2.0 applications, and their implications for data publishing, sharing, and preservation*.

In this sense, there have already been proposals to use a Web 2.0-based Scientific Social Community (SSC) model for the enhancement of the bioinformatics research (e.g. Zhang *et al.* 2009). According to these models even scientific databases should be constructed in a way that lets users over the Internet have access, modify and incorporate their own data into main databases. In theory this would be ideal to solve some of the problems of scientific database location and heterogeneity of data. However, it requires, as it was observed in a previous section, a huge collaborative effort from the major research institutions and researchers worldwide. An example of this type of technology applied to biological databases include mobile devices with basic biological information about different species within a particular group that can be used in the field for identification purposes. This is very similar to what is used in several museums worldwide for users to access information on art pieces.

Given the different models of data sharing, the advantages of the Internet and the type of technology associated to the Web 2.0 it would seem easy that all the biological data available could be put together into an unique multinational database. However, as Moritz (2004b) has noted, the universe of data, information and knowledge available thus far by the scientific community is weakly integrated and any efforts of integration would be immediately confronted by a complex array of legal and cultural barriers. With this situation facing the future of scientific data release in a world of changes, it seems necessary that we as scientists try to find an agreement that validates our social responsibility without compromising the efforts of our daily work. This is not an easy task and will probably require, besides scientists, a highly international group of competent professionals from areas such as international law, sociology and informatics. We know some efforts have been carried out in recent years; hopefully they have already established a milestone for future approaches.

Conclusions & Recommendations

Tropical ecosystems will be both influenced by and influence patterns of global change. Global change is by definition a global issue and requires a concerted global effort in order to understand and mitigate the effects on natural systems. While scientific progress in the past has been largely due to individuals or small groups working on specific questions, questions about global change require a integrated approach. What is needed is a *globalization of global change research*. In this paper, we have identified many of the issues that hinder progress in this field. In summary, we make the following recommendations:

- Shifting of funding to longer term and more inclusive, to take advantage of local knowledge in tropical countries
- Investigators in all parts of the world should have access to a global pool of research funds; and should be strongly encouraged, if not required, to collaborate.
- Data collection should be standardized as much as possible between experiments, although with allowances for alternate methods and the incorporation of technological advances. When changes in sampling protocol are justified, researchers should strive to collect data using both methods (new and old) for a transition period to calibrate the measurements.
- There are already several long-term monitoring protocols freely available on the

internet. These methods have been tried, tested and improved over several years under different conditions. Establishment of new protocols for the acquisition of the same or similar data would be a waste of time and funds. Ample dissemination of them is thus highly encouraged.

- If there is a need to develop new protocols, researchers must ensure the compatibility and documentation of their studies, as well as take into account the diversity of ecosystems, countries and cultures in the tropics. This, in order to ensure, that at any point, the establishment of additional experimental sites will be viable and available widely across disciplines as well as by the decision-making community (Sanchez-Azofeifa *et al.* 2005).
- Long-term monitoring and data collection should be acknowledged as a worthy contribution to science. Ideally the scientific infrastructure (tenure, promotion, access to grant funds) should reward scientists for data collection in addition to analysis and theoretical work.
- Global change research in the tropics will also benefit from willingness (or requirements) to make data available in a timely manner.
- More sincere effort needs to taken on the part of researchers to compile and archive complete metadata, data, and analytical processes.
- Field stations should place a high priority on archiving data (with detailed metadata) or require it from researchers at the conclusion of projects.

References

Baker T, Phillips OL, Malhi Y, Almeida S, Arroyo L, Di Fiore A, Erwin T, Higuchi N, Killeen TJ, Laurance SG, Laurance WF, Lewis SL, Monteagudo A, Neill DA, Vargas PN, Pitman NCA, Silva NM, Martínez RV (2004) Increasing biomass in Amazonian forest plots. *Philosophical Transactions of the Royal Society, Series B* 359: 353-365.

Chamber, JQ, GP Asner, DC Morton, LO Anderson, SS Saatchi, FDB Espirito-Santo, M Palace and C Souza (2007) Regional ecosystem structure and function: ecological insights from remote sensing of tropical forests. *Trends in ecology and Evolution* 22: 414-423.

Chave JR, Condit S, Lao JP, Caspersen RB, Foster, Hubbell SP (2003) Spatial and temporal variation in biomass of a tropical forest: results from a large census plot in Panama. *Journal of Ecology* 91: 240-252.

Chave JR, Condit HC, Muller-Landau SC, Thomas PS, Ashton S, Bunyavejchewin LL, Co HS, Dattaraja SJ, Davies S, Esufali CEN, Ewango KJ, Feeley RB, Foster N, Gunatilleke S, Gunatilleke P, Hall TB, Hart C, Hernández SP, Hubbell A, Itoh S, Kiratiprayoon JV, LaFrankie S, Loo de Lao JR, Makana Md, Nur Supardi Noor A, Rahman Kassim C, Samper R, Sukumar HS, Suresh S, Tan J, Thompson Ma DC, Tongco R, Valencia M, Vallejo G, Villa T, Yamakura JK, Zimmerman, Losos EC (2008) Assessing Evidence for a Pervasive Alteration in Tropical Tree Communities. *PLoS Biol*ogy 6: e45.

Clark DA, Piper SC, Keeling CD, Clark DB (2003) Tropical rain forest tree growth and atmospheric carbon dynamics linked to interannual temperature variation during 1984-2000. *Proceedings of the National Academy of Sciences* 100: 5852-5857.

Clark DA (2004) Sources or sinks? The responses of tropical forests to current and future climate and atmospheric composition. *Philosophical Transactions of the Royal Society of London Series B* 359: 477-491.

Clark DB, Castro CS, Alvarado LDA, Read JM (2004) Quantifying mortality of tropical rain forest trees using high-spatial-resolution satellite data. *Ecology Letters* 7: 52-59.

Clark DA (2007) Detecting tropical forests' responses to global climatic and atmospheric change: Current challenges and a way forward. *Biotropica* 39: 4-19.

Condit R (1995) Research in large, long-term tropical forest plots. *Trends in Ecology and Evolution* 10: 18-22.

Colwell RK, Brehm G, Cardelus CL, Gilman AC, Longino JT (2008) Global warming, elevational range shifts, and lowland biotic attrition in the wet tropics. *Science* 322: 258-261.

Conley W, Brunt JW (1991) An institute for theoretical ecology? Part V: Practical data management for cross-site analysis and synthesis of ecological information. *Coenoses*. 6: 173-180.

DiGiusto G (2005) Facts Are Stubborn Things, but Should the Law Protect Them? The Non-Regime for Database Protection. Proceedings of the Annual Meeting of the American Political Science Association.

Druce DJ, Shannon G, Page BR, Grant R, Slotow R (2008) Ecological Thresholds in the Savanna Landscape: Developing a Protocol for Monitoring the Change in Composition and Utilisation of Large Trees. *PLoS ONE* 3: e3979.

Ferraz G, Marinelli CE, Lovejoy TE (2009) Biological monitoring in the Amazon: recent progress and future needs. *Biotropica* 40: 7-10.

Gantz JF (2007) The Expanding Digital Universe: A Forecast of Worldwide Information Growth through 2010. I.D.C. Whitepaper.

Janzen DH (2004) Now is the time. *Philosophical Transactions of the Royal Society of London, Series B*. 359: 731-732.

Jones MB, Schildhauer MP, Reichman OJ, Bowers S (2006) The New Bioinformatics: Integrating Ecological Data from the Gene to the Biosphere. *Annual Review of Ecology, Evolution and Systematics* 37: 519-544.

King DA 2004 The scientific impact of nations: What different countries get for their research spending. *Nature* 430: 311-316.

Laurance WF, Oliveira AA, Laurance SG, Condit R, Nascimiento HEM, Sanchez-Thorin AC, Lovejoy TE, Andrade A, D'Angelo S, Ribeiro JE, Dick CD (2004) Pervasive alteration of tree communities in undisturbed Amazonian forests. *Nature* 428: 171-175.

Leimu R, Koricheva J (2005) Does scientific collaboration increase the impact of ecological articles? *Bioscience* 55: 438-443.

Lewis SL, Malhi Y, Phillips OL (2004) Fingerprinting the impacts of global change on tropical forests. *Philosophical Transactions of the Royal Society, Series B* 359: 437-462.

Lovett GM, Burns DA, Driscoll CT, Jenkins JC, Mitchell MJ, Rustad L, Shanley JB, Likens GE, Haeuber R (2007) Who needs environmental monitoring? *Frontiers in Ecology and the Environment* 5: 253-260.

Michener WK, Brunt JW, Helly JJ, Kirchner TB, Stafford SG (1997) Nongeospatial metadata for the ecological sciences. *Ecological Applications*. 7: 330-342.

Moritz TD (2004a) Building the Biodiversity Commons. *D-Lib Magazine*. DOI 10.1045.

Moritz TD (2004b) "Conservation Partnerships in the Commons? Sharing data and information, experience and knowledge, as the essence of partnerships", Museum International.

NERC: Natural Environmental Research Council. (Page published on 29 July, 2008) Ecosystem services for poverty alleviation. www.nerc.ac.uk/research/programmes/espa. Polaris House, North Star Avenue, Swindon, SN2 1EU, United Kingdom (Page viewed on 11 May, 2009).

NSF: National Science Foundation. (Page published on April 29, 2009) National Science

Foundation: Where discoveries begin. www.nsf.gov. 4201 Wilson Boulevard, Arlington, Virginia 22230, USA (Page viewed on 11 May, 2009).

Parmesan C (2006) Ecological and Evolutionary Responses to Recent Climate Change. *Annual Review of Ecology, Evolution, and Systematics* 37: 637-669.

Phillips, O., T. Baker, T. Feldpausch and R. Brienen (2008) RAINFOR field manual for plot establishment and remeasurement. Document available at http://www.geog.leeds.ac.uk/projects/rainfor/manuals/RAINFOR%20field%20manual%20version%202008%20ENG.pdf (accessed 12 May 2009).

Portner HO, Farrell AP, Knust R, Lannig G, Mark FC, Storch D (2009) Adapting to climate change response. *Science* 323: 876-877.

Pounds JA, Bustamante MR, Coloma LA, Consuegra JA, Fogden MPL, Foster PN, La Marca E, Masters KL, Merino-Viteri A, Puschendorf R, Ron SR, Sanchez-Azofeifa GA, Still CJ, Young BE (2006) Widespread amphibian extinctions from epidemic disease driven by global warming. *Nature* 439: 161-167.

Pounds JA, Fogden MPL, Campbell JH (1999) Biological response to climate change on a tropical mountain. *Nature* 398: 611-615.

Sánchez-Azofeifa GA, Quesada M, Rodríguez JP, Nassar JM, Stoner KE, Castillo A, Garvin T, Zent EL, Calvo-Alvarado JC, Kalacska M, Fajardo L, Gamon JA, Cuevas-Reyes P (2005) Research Priorities for Neotropical Dry Forests. *Biotropica* 37: 477-485.

Svenning JC, Condit R (2008) Biodiversity in a warmer world. *Science* 322: 206-207.

Spellerberg IF (2005) Monitoring ecological change. 2nd Edition. Cambridge University Press, Cambridge.

Trembly A (2009) Internet Regulation: Distasteful but Inevitable. *National Underwriter Life & Health* 113: 14.

Wright SJ (2005) Tropical Forests in a Changing Environment. *Trends in Ecology & Evolution* 20: 553-560.

Whitfield SM, Bell KE, Philippi T, Sasa M, Bolanos F, Chaves G, Savage JM, Donnelly MA (2007) Amphibian and reptile declines over 35 years at La Selva, Costa Rica. *Proceedings of the National Academy of Sciences* 104: 8352-8356.

Williams SE, Shoo LP, Isaac JL, Hoffmann AA, Langham G (2008) Towards an integrated framework for assessing the vulnerability of species to climate change. *Plos Biology* 6: 2621-2626.

Young BE, Lips KR, Reaser JK, Ibanez R, Salas AW, Cedeno JR, Coloma LA, Ron S, La Marca E, Meyer JR, Munoz A, Bolanos F, Chaves G, Romo D (2001) Population declines and priorities for amphibian conservation in Latin America. *Conservation Biology* 15: 1213-1223.

Zhang Z, Cheung KH, Townsend JP (2009) Bringing Web 2.0 to Bioinformatics. *Briefings in Bioinformatics* 10: 1-10.

Appendix 1: Survey for course participants

Global Change Research and Data Management in Tropical Ecosystems

1. Who should pay the most for long-term global change research in tropical ecosystems?
   a. Tax payers in temperate, developed countries
   b. Tax payers in tropical, developing countries
   c. Nongovernmental organizations in temperate, developed countries
   d. Nongovernmental organizations in tropical, developing countries
   e. There should be a global pool of money from many governments and organizations

2. Regarding allocation of funds for long-term global change research in tropical ecosystems, please rank the following from least (1) to most (4) important, or choose e.
   a. _____Citizenship of investigator
   b. _____Project description and aims
   c. _____Investigator's prior experience, as indicated on his/her CV
   d. _____Resources available to investigator at home institution (e.g., instruments)
   e. All of these are equally important

3. Collaboration between investigators in temperate, developed countries and tropical, developing countries
   a. Is ideal and should be required
   b. Can be good, but should not be required
   c. Is not good and should not be required

4. Renewable grants that support long-term global change studies in tropical ecosystems should last
   a. < 3 years
   b. 3 – 5 years
   c. > 5 years

5. Renewable grants that support research stations in tropical ecosystems should last
   a. < 3 years
   b. 3 – 5 years
   c. > 5 years

6. Sampling methods for long-term global change research in tropical ecosystems should be
   a. More standardized, e.g., to allow comparisons across broad spatiotemporal scales
   b. Less standardized, e.g., to encourage development of new and improved sampling methods

7. How should data be stored?
   a. Locally by the investigator(s)
   b. Nationally by the government
   c. Globally by a multi-national body
   d. Other ___________________________

8. Please check any of the following metadata and data collections that you are familiar with:
   ☐ Knowledge Network for Biocomplexity
   ☐ NBII Metadata Clearinghouse
   ☐ NSDI Metadata Clearinghouse
   ☐ GBIF Taxonomic Collections
   ☐ TOPP
   ☐ Kruger National Park/SAEON
   ☐ TreeBase
   ☐ Storage Resource Broker
   ☐ ORNL DAAC
   ☐ Global Change Master Directory
   ☐ ESA Data Registry
   ☐ Ecological Archives
   ☐ Ocean Biological Information Systems
   ☐ Global Population Dynamics Database
   ☐ VegBank
   ☐ Open-Source Project for a Network Data Access Protocol (OPeNDAP)/Integrated Ocean Observing System (IOOS)

9. How should data be shared?
   a. Investigators should not be required to share data
   b. Data should be posted on the internet
   c. Data should be published (e.g., Ecological Archives)
   d. Other ___________________________

10. What is one way that you would like to see global change research or data management in tropical ecosystems improved?