4. Monitoring and Evaluation of Changes in Carbon Stock

In Figure 5, we present an overview of the approach used in LBNL's MERVC guidelines for evaluating changes in the carbon stock. During the monitoring and evaluation stage, gross changes in the carbon stock are measured, using one or more of the following monitoring and evaluation methods: modeling, remote sensing, and field/site measurements (Section 4.2). The baseline is also re-estimated, accounting for free riders (Section 4.3). The net change in the carbon stock is equal to the gross change in the carbon stock minus the re-estimated baseline.

During the implementation of the project, monitoring of project activities is conducted periodically to ensure the project is performing as designed. We expect most, if not all, of the above activities to be performed by project developers and their contractors.[1] While the project is being implemented, however, we expect periodic (e.g., annual) reviews by third-party verifiers (to avoid conflicts of interest), leading to certification (see Sections 6 and 7). These verifiers might be the same independent reviewers who assessed the project proposal at the registration stage (personal communication from Johannes Heister, The World Bank, Jan. 12, 1999). As noted in Section 6, verification of carbon sequestration would be performed at certain intervals during the time the project is scheduled to sequester carbon.

This section introduces some of the basic data collection and analysis methods used to estimate changes in the carbon stock and associated impacts. The methods vary in cost, accuracy, simplicity and technical expertise required. Tradeoffs will need to be made for choosing the appropriate methods: e.g., level of accuracy and cost of data collection.

[1] An alternative approach is to require only certified professionals to conduct the monitoring and evaluation, as required when institutions of higher education enter into energy performance-based contracts in Texas (Texas Higher Education Coordinating Board 1998).

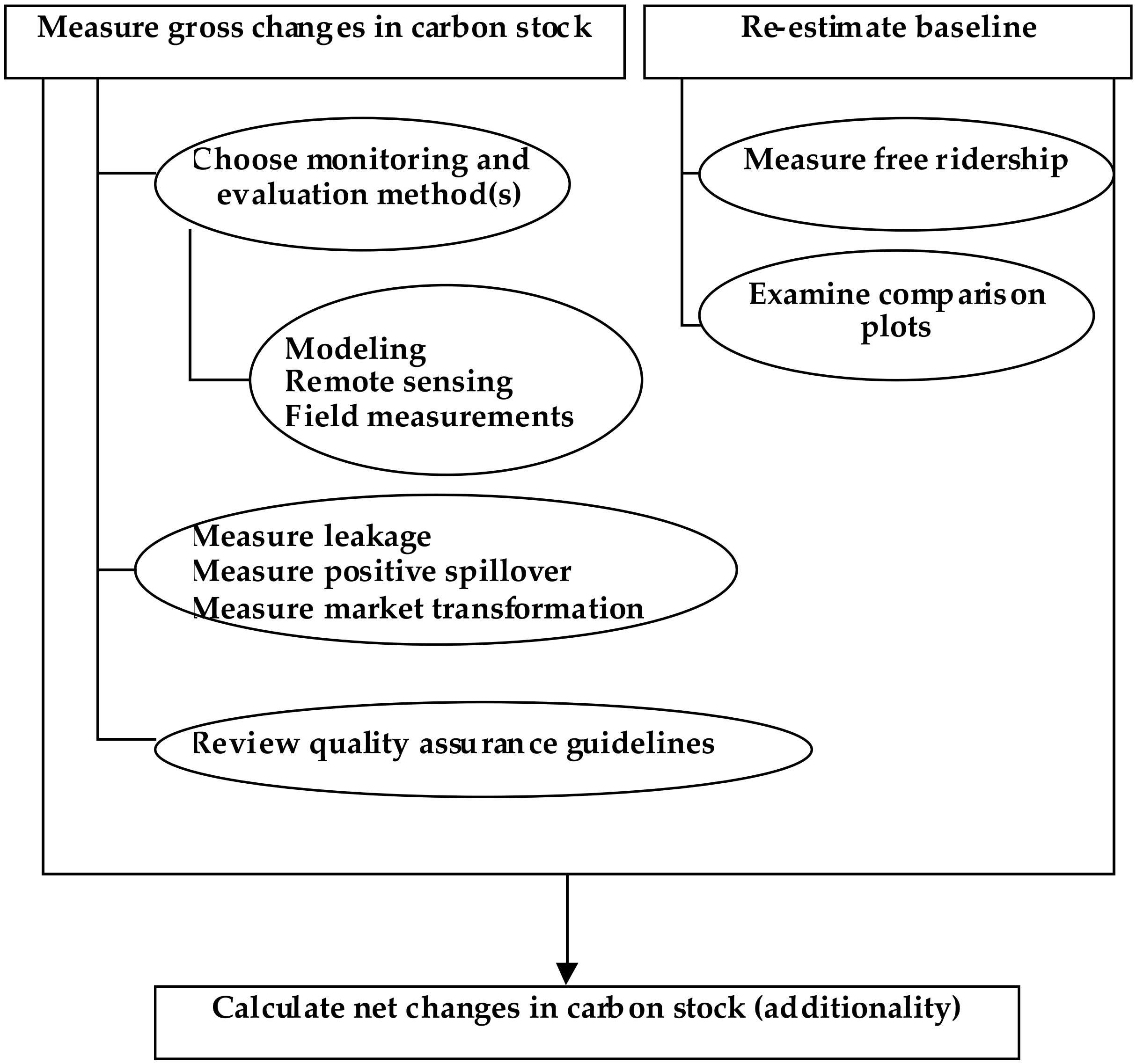

Fig. 5. Evaluation Overview

4.1. Methodological Issues

Prior to reviewing the data collection and analysis methods used for measuring changes in the carbon stock and associated impacts, we first discuss two key methodological issues: measurement uncertainty, and the frequency and duration of monitoring and evaluation. As indicated elsewhere, these issues are not only addressed in the monitoring and evaluation stage but should also be examined in the project design stage.

4.1.1. Measurement uncertainty

While there are several types of uncertainty that can affect the actual realization of GHG reductions, uncertainty in the measurement of GHG reductions needs to be accounted for when presenting monitored and evaluated data.[1] Measurement uncertainties include the following: (1) the use of simplified representations with averaged values (especially emission factors); (2) the uncertainty in the scientific understanding of the basic processes leading to carbon sequestration, and emissions and removals of non-CO_2 GHG;[2] and (3) the uncertainty in measuring items that cannot be directly measured (e.g., project baselines). Some of these uncertainties vary widely by type of project (depending on approach, level of detail, use of default data or project specific data, etc.), and length of project (e.g., short-term versus long-term). It is important to provide as thorough an understanding as possible of the uncertainties involved when monitoring and evaluating the impacts of forestry projects.

As an example, in the assessment of the project design of Costa Rica's Protected Areas Project (PAP), the risk and uncertainty assessment identified and quantified the carbon offsets that had any type of risk or uncertainty associated with them (SGS 1998). These were then set into a buffer to provide insurance for the carbon offsets that were authorized for sale (i.e., outside the buffer). The principal source of uncertainty in these estimates was the rate of deforestation occurring outside the protected areas (see Busch et al. 1999).

Because of the difficulties and uncertainties in estimating changes in carbon stock, the level of precision and confidence levels associated with the changes need to be identified.[3] Project

[1] Other types of uncertainty include the following: (1) project development and construction uncertainty, i.e., the project won't be implemented on time or at all, even though funds have been spent on project development; (2) operations and performance uncertainty (e.g., if forest management practices are not implemented as projected, or if the project is not maintained for a reasonably long time, then net carbon sequestered will change); and (3) environmental uncertainty (IPCC 1996; USAID 1996; UNFCCC 1998a). In the case of environmental uncertainty, some sites may have a higher chance of disturbances, such as wildfire ignitions, flood overflows, avalanches, etc. Typically, fire-related species are established in places where fire is much more of a risk than in other places; similarly, riparian forests are re-established in areas where flood flows in the first few years of a project are likely to destroy the effort. Hence, project developers should provide a description of the project developer's experience, existing warranties, the performance history of previous projects, and plans to reduce uncertainty. The political and social conditions that exist that could potentially affect the credibility of the implementing organizations (e.g., political context, stability of parties involved and their interests, and potential barriers) also need to be described.

[2] For example, many knowledge gaps still remain in our understanding of the fundamental mechanisms responsible for soil carbon dynamics (Lal et al. 1997).

[3] Unless otherwise noted, we assume normal distributions, represented by a normal, bell-shaped curve in which the mean, median and mode all coincide.

developers and evaluators should report the precision of their measurements and results in one of two ways: (1) quantitatively, by specifying the standard deviation around the mean for a bell-shaped distribution, or providing confidence intervals around mean estimates; or (2) qualitatively, by indicating the general level of precision of the measurement (e.g., low, medium or high).

It is unclear at this time on how uncertainty will be treated in the calculation and crediting of reduced carbon emissions and increased carbon sequestration. At a minimum, the most conservative figures should be used at every stage of calculation (e.g., the lower boundary of a confidence interval). The qualitative assessment of uncertainty is more problematic, however, some type of discounting or debiting could be used to adjust the amount of carbon sequestered in situations where there is a great deal of uncertainty. Where there is substantial uncertainty, project developers need to design higher quality forestry projects so that impacts are more certain.

In conclusion, the evaluation of forestry projects should: (1) evaluate the project's contingency plan, where available, that identifies potential project uncertainties and discusses the measures provided within the project to manage the uncertainties; (2) identify and discuss key uncertainties affecting all emission estimates; (3) assess the possibility of local or regional political and economic instability and how this may affect project performance; and (4) provide confidence intervals around mean estimates.

4.1.2. Frequency and duration of monitoring and evaluation

The frequency of monitoring and evaluation will most likely be linked to the schedule of transfer of carbon credits. It is possible that these credits could be issued on an annual basis.[1] The monitoring period may last longer than the project implementation period: for example, if it takes three years to complete a reforestation project, net carbon sequestered from the project will continue beyond the three years. Moreover, after the crediting of the emissions reductions is over, a system is still needed for keeping track of the carbon from these projects. Hence, the sustainability of forestry projects (excluding substitution projects, such as bioenergy plantations) is critical if the impacts from these projects are to persist. Therefore, information is needed on the institutional capabilities and support for implementing the project over the project's lifetime and on the uncertainties of a project (see Section 4.1.1).

[1] Other models are possible (e.g., up-front lump-sum payment), but unlikely since the issuance of certified emission reductions occurs after a verification process.

The institutional, community, technical and contractual conditions likely to encourage sustainability of forestry projects are of utmost concern. In some cases, encouraging the participation of community members in the development and implementation of forestry projects will help to ensure the longevity of a project, although the design and implementation process may take longer and costs will increase. Sustainability will also increase by encouraging operations and maintenance, providing spare parts and equipment, and making sure technical expertise is available. Finally, contracts can incorporate provisions that lead to debiting of emission reduction units (for the host and/or investor country) if a project does not last as long as expected.

The frequency of monitoring and evaluation will also depend on the carbon pools being affected by the project (Section 2.1). Each carbon pool has a different rate of change: e.g., above-ground biomass often experiences greater rates of change in carbon than soils. For those carbon pools that undergo relatively rapid changes in carbon (e.g., above-ground biomass, for fast-growing species), monitoring of the carbon pools should be done on an annual basis. As discussed in Section 4.2.5, sampling will help to reduce these costs. However, monitoring should also account for the type of project being implemented. For example, for afforestation projects, monitoring and evaluation could be done less frequently in the first few years of the project, since significant changes in carbon will not be occurring (in fact, the standard error of the estimates may be larger than the actual growth). After five years, annual monitoring may be warranted.

The monitoring of soil carbon may not need to be conducted on an annual basis, for two reasons. First, the monitoring of soil carbon is relatively more expensive. Second, in undisturbed areas, soil carbon does not change dramatically from year to year. Hence, soils could be visually inspected annually to note the absence or presence of soil-altering events, and detailed monitoring could be done every five years. However, for areas experiencing severe disturbances (or continuous soil disturbances, like farming), monitoring should be conducted more frequently, perhaps annually. In addition, because organic carbon levels in soils go up and down due to seasonal variation during the year (Lal et al. 1998), monitoring of soil carbon should be conducted consistently: i.e., at the same time each year. If sampling is not done consistently, then the temporal fluctuations may overwhelm the real changes in carbon. While seasonal changes in total soil organic matter are hard to detect, some components of soil organic matter (e.g., dissolved organic carbon or the biomass carbon) may change among seasons (personal communication from Rattan Lal, Ohio State University, Jan. 30, 1999). On the other hand, significant changes in total soil organic matter may be only detectable over a 2 to 3 year period, with longer periods more desirable.[1] Ideally, measurement of soil carbon levels should be done when the soil carbon levels are at their average levels for both reference and project cases.

In addition to measurements, records should be kept on disturbances at the sites, whether human-made (e.g., thinnings) or natural (e.g., pest infestation). For forest products, guidance for the frequency of monitoring is more difficult to provide, since the removal of (i.e., demand for) wood products is a function of socioeconomic pressures, as well as natural replacement. As a conservative estimate, we encourage annual monitoring of wood products.

Finally, where more than one project is being implemented, evaluators should evaluate a project by its permanence or lack of permanence — this will be reflected in "project lifetime," which may be different than an expected lifetime of a project as initially proposed by developers. The project lifetime is a function of the type of carbon pool affected (e.g., soils versus above-ground woody biomass) and the probability of an occurrence of a natural or human-made disturbance (e.g., fires). For example, if a project area is likely to undergo serious changes after 20 years, then the changes in the carbon stock for that project are limited to that 20-year lifetime. The value of those changes may be less than for changes in the carbon stock from similar projects with longer project lifetimes (e.g., 30 years). Accompanying the evaluation, the evaluator should provide a list of indices that demonstrate the potential for permanence: e.g., type and number of income groups targeted by project, potential socioeconomic impacts addressed (see Section 8.2), potential sources of uncertainty addressed (see Section 4.1.1), etc.

4.2. Measurement of Gross Changes in Carbon Stock

These guidelines are to be used to measure the changes in carbon stocks as accurately as is practical, accounting for all positive emissions of carbon from forests (e.g., from the combustion and decay of organic matter and the use of fossil fuels in machinery) and negative emissions (capture) of carbon through photosynthesis in forests. Since forestry activities typically trigger a sequence of effects that change through time, the measurement of changes in carbon stocks must account for these dynamic effects (e.g., from the time a forest is established until a forest is removed by harvest or a natural disturbance).

[1] For example, after sufficient time (e.g., 5 to 10 years), statistically, significant differences in soil organic carbon have been observed in natural, inadvertent, and planned experiments (e.g., Izaurralde et al. 1998; Paul et al. 1997; Smith et al. 1997).

The measurement of a project's carbon fixation necessitates specialized tools and methods drawn largely from experience with forest inventories and ecological research. Monitoring and verifying carbon accumulation in forestry projects must be cost effective and accurate. Monitoring systems should be built upon standard forestry approaches to biomass measurement and analysis, and apply commonly accepted principles of forest inventory, soil science and ecological surveys. Specific methods and procedures should be assembled on a project-specific basis, with the types and extent of monitoring ultimately determined by the relative costs and quantity of carbon return by each measurement type.

Three general monitoring techniques can be used to monitor carbon fixed through forestry projects (based on MacDicken 1997): (1) modeling, (2) remote sensing, and (3) field/site measurements, including biomass surveys (which includes research studies; surveys; the monitoring of wood production and end products; and forest inventories) and destructive sampling. Many of these techniques can be used together.

4.2.1. Establishing the monitoring domain

Different techniques are available for assessing multiple monitoring domains in forestry projects (Andrasko 1997). At the national scale, remote sensing can be used to detect land-use and land-cover changes. At the regional scale, remote sensing can be used with ground-truthing and forest inventory techniques. And at the project level, remote sensing, ground-truthing, creation of permanent plots, forest inventory data or surveys, or allometric techniques can be used (see Section 3.2 and Box 2).

Currently, there are weak linkages in assessing multiple monitoring domains (Andrasko 1997). One potential solution to strengthening these linkages is the use of "nested monitoring systems" where an individual project's monitoring domain is defined to capture the most significant GHG fluxes and where provisions are made for monitoring carbon stocks and GHG flows outside of the project area by regional systems or national GHG inventory monitoring systems (Andrasko 1997).

4.2.2. Modeling

Modeling the impacts of certain forestry practices on carbon flows into and out of forest carbon sinks can be used for estimating annual flows of carbon. The models are used to predict future carbon flows, but they do not measure the actual changes. The modeled estimates of carbon storage over time must be checked using one of the techniques described below (i.e., remote sensing with ground truthing or field/site measurement).

Models start from an estimate of a carbon stock for a specific forest type at a specific site. Then, based on information from forest practices, the models develop estimates of annual carbon flows. This approach relies on a series of highly simplified assumptions to estimate total carbon sequestration. For example, assumptions may include: the number of trees planted in either woodlots or agroforestry systems, initial stocking rates, mean annual stemwood volume increments, a biomass multiplier factor, and harvest rates. The assumptions are then inputted into a model to estimate the amount of sequestered carbon. The models need to be corrected/calibrated with measured data periodically as well as with other approaches. For example, approaches that estimate forest productivity by timber volume may be compared with other approaches, such as allometrically derived carbon estimates that incorporate relationships between tree or stand physiological parameters (e.g., diameter, height, weight, tapper (the change in diameter over height) and carbon content (Box 2) (Hamburg et al. 1997; Schroeder et al. 1997; Brown 1997). The accuracy of these methods will depend on many factors, including the precision of the equations and the homogeneity of the forest (e.g., allometric equations are simpler and more accurate for homogeneous forests and more complex and less accurate for heterogeneous forests).

Some models are already available for simple conditions and standard treatments, such as tree planting on agricultural land. The Land Use and Carbon Sequestration (LUCS) model is a project-based computer model that tracks the changes in carbon density associated with land use changes (e.g., conversion of forested areas to agriculture) (Faeth et al. 1994; MacDicken 1998).[1] Direct measurements and default assumptions are used to calculate the changes and impacts. The LUCS model has been used in evaluating an agroforestry project on marginal hillsides in Guatemala (Trexler et. al. 1992).

[1] The LUCS model is available from the Lawrence Berkeley National Laboratory.

Box 2

Modeling Example

To reduce the costs of carrying out forest mitigation projects, there is an incentive to use volume-based estimates of carbon content of forested lands and existing allometric equations developed outside the project region to estimate forest carbon. The comparison of generalized allometric equations with volumetric estimates was conducted in a study of Russian forests. In addition to comparing volume and allometrically derived carbon estimates, the applicability of using generalized allometric equations was evaluated by comparing estimates generated utilizing North American allometric equations with locally derived Russian equations.

Evaluation method: The volume and allometrically derived carbon estimates of 51 Russian forests were compared. Russia provided an ideal setting for comparing these two approaches since large data sets, collected from across the country, included tree volume, tree weights, and stand characteristics. Representative stand data were selected from information in the Russian National Forest Inventory for forests with species compositions representative of the dominant vegetation of the two regions of interest. A system of phytomass/volume ratios was developed to convert timber volume to stand carbon. Construction of the allometric equations utilized data from individual trees and shrubs collected in the regions. One thousand individual trees of the five dominant species found in the regions were destructively sampled: all trees were cut and divided into four parts for development of the allometric equations: stem, branches, leaves, and roots. Allometric equations were developed for the trees of interest. On each plot, the heights of 10-12 trees were measured and allometric equations relating tree height to diameter at breast height were developed and carbon contents were assumed. Volume-based estimates of carbon content of forest stands involved the application of zonal and regional species phytomass/volume ratios, evaluated using the forest phytomass and productivity database available in Russia. The phytomass/volume ratios utilized the same carbon/dry weight percentages as were used in the allometric equations.

Findings: Volumetrically and allometrically derived carbon estimates of 51 Russian forests were very similar. The error associated with volumetrically derived carbon estimates varied with species composition. For some species, there was no apparent difference between volumetric and allometric estimates, but for others it averaged 15%. The results also suggest that it is appropriate to utilize allometric equations developed for one species for estimating the carbon content of another species growing in a different region, as long as they are phenotypically similar. Both volumetric and allometric approaches for estimating carbon are useful. For regional based studies of forest carbon, volumetric approaches are preferred because they are easy to use. For stand-based estimates of forest, carbon allometric approaches provide greater reliability.

Sources: (1) Hamburg, S., D. Zamolodchikov, G. Korovin, V. Nefedjev, A. Utkin, J. Gulbe, and T. Gulbe. 1997. "Estimating the Carbon Content of Russian Forests: A Comparison of Phytomass/Volume and Allometric Projections," *Mitigation and Adaptation Strategies for Global Change* 2(2-3): 247-265; (2) Schroeder, P., S. Brown, J. Mo, R. Birdsey and C. Cieszewski. 1997. "Biomass Estimation for Temperate Broadleaf Forests of the United States Using Inventory Data," *Science* 43(3): 424-434; (3) Brown, S., 1997. *Estimating Biomass and Biomass Change of Tropical Forests: A Primer*. FAO Forestry Paper 134, Food and Agricultural Organization of the United Nations, Rome, Italy.

Soil organic matter and ecosystem models play an important role in understanding land management and soil organic carbon sequestration relationships and for projecting changes in soil organic carbon through time (Parton et al. 1995; Smith et al. 1997). The rate of soil organic carbon decomposition is usually well represented as a first-order process where the amount converted to CO_2 per unit time depends on the current size of the various soil organic carbon fractions times their rate constants (Smith et al. 1997). Since the amounts present in each carbon fraction depends on management history, these amounts must be accurately accounted if the model estimates of soil organic carbon dynamics are to be realistic. Generally, information on previous management history is less complete than needed to establish adequate initial conditions for models. When management history is well known for a period of at least 20-50 years, many soil organic carbon models do well in simulating management-induced soil organic carbon changes (Smith et al. 1997). Model validation remains an important step for validating models assumptions.

The Graz/Oak Ridge Carbon Accounting Model (GORCAM) is another model that can be used to examine the impact of forestry projects on carbon emissions (Schlamadinger and Marland 1996). GORCAM provides a simplified description of carbon stocks and flows associated with the management of forests. GORCAM calculates carbon accumulation in plants, in short- and long-lived wood products, in fossil fuels not burned because biofuels are used instead, and in fossil fuels not burned because production and use of wood products requires less energy than does production and use of alternative materials that provide the same service (Marland et al. 1997). GORCAM has been used to evaluate the impact on carbon emissions by biofuel district heating systems being installed or proposed in Vermont (McLain 1998), as well as estimating the amount of carbon sequestered by a sustainable forestry management project in Mexico (Bird et al. 1997).

More complex but promising models are being developed (USDOE 1994). Simple modeling requires relatively little time and effort, however, the gross estimates are probably neither accurate nor precise (MacDicken 1997). In general, field/site measurements are preferred over standard tables and computer models, because site-specific field studies provide higher quality data and thus higher credibility, although at a higher cost.

4.2.3. Remote sensing

Remote sensing (along with ground-based measurements) can be used to monitor land area changes, map vegetation types, delineate strata for sampling, and assess leakage and base case assumptions (Box 3). Remote sensing is defined as the acquisition of data about an object or scene by a sensor that is far from the object (Colwell 1983; see also Slater 1980; Swain and Davis 1978; Wilkie and Finn 1996). Aerial photography, satellite imagery, and radar are all forms of remotely sensed data. Usually, remote sensing refers to the following two types: (1) “high-level” remote sensing that uses satellite imagery, and (2) “low-level” remote sensing that relies on aerial photography.

Box 3

Remote Sensing Example

The need for a statistical approach to sampling remote sensing databases is crucial for monitoring deforestation and other more comprehensive land use/land-cover processes. Even though a great deal of information exists at the project level, there is still much uncertainty when there is a need to scale up from the project scale to regional or national scales. In addition to sampling, additional aspects such as sensor spatial and spectral resolution, frequency of acquisition of remote sensing information, and economic costs are key components of the monitoring program and its methodological development.

Evaluation method: Sanchez-Azofeifa et al. (1997) used the Landsat tile system (World Reference System 2) as a sampling frame for the selection of remote sensing data. A remotely sensed data set of a wall-to-wall assessment of deforestation in the Amazon basin was used. The data base consisted of land cover change information extracted from 228 satellite scenes for 1978 and 1988. Deforestation, primary forest, clouds, and naturally occurring non-forest were the main topological attributes. A stratified population was used prior to selecting a sample for random sampling. The population was stratified first by eliminating all scenes with an area of more than 30% of non-natural forest area and then stratified by total deforestation, deforestation rate, and the permanence of deforestation.

Evaluation concerns: Because of the patchiness of deforestation, random sampling of Landsat scenes can produce significant errors when the goal is to estimate total deforestation.

Findings: Stratification based on permanence contributed to the reduction of error in the estimation of total deforestation when contrasted to random sampling without stratification. Random sampling has the potential for extreme over- or under-estimation of total deforestation. Reductions in error were achieved only when very high sampling densities were obtained. When a new level of stratification was applied, very accurate estimates of the total area deforested were obtained using low sample densities.

FAO Study: In the second phase of the 1990 Forest Resources Assessment, the Food and Agricultural Organization (FAO) of the United Nations employed a statistical survey using remote sensing to sample the forest cover in tropical forests. The same analyst based the approach for the second phase on the comparison of satellite imagery from two dates at the same time, using a uniform classification throughout the tropics. By using this approach, class to class changes in land cover (e.g., from grassland to forest, or vice versa) could be detected and depicted in change matrices according to regions and climatic zones. Information on class to class changes is new and adds substantially to the understanding of the processes of vegetation and deforestation.

Sources: (1) Sanchez-Azofeifa, G., D. Skole, and W. Chomentowski. 1997. "Sampling Global Deforestation Databases: The Role of Persistence," *Mitigation and Adaptation Strategies for Global Change*, 2(2-3):177-189. (2) Food and Agricultural Organization (FAO) of the United Nations. 1996. *Forest Resources Assessment 1990: Survey of Tropical Forest Cover and Study of Change Processes*. FAO Forestry Paper 130, Food and Agricultural Organization of the United Nations, Rome, Italy.

High-level remote sensing. Many national and international projects and programs have made use of remote sensing with satellites for land cover change research at a national or international level (FAO 1996; Skole et al. 1997). This type of remote sensing can be done every 5-10 years, in combination with low-level remote sensing. The Face Foundation in the Netherlands and Winrock International have used satellite imagery for evaluating forestry projects (Face Foundation 1997;

MacDicken 1998).[1] Remote sensing has been used by several researchers in measuring deforestation in tropical forests in Central and South America (e.g., Dale et al. 1994; Sanchez-Azofeifa et al. 1997; Sanchez-Azofeifa and Quesada-Mateo 1995; Skole and Tucker 1993; Stone et al. 1991). Attempts to estimate biomass from remote sensors have generally been costly and have had mixed results (MacDicken 1997). To date, no one has measured carbon using remote sensing (Brown 1996; MacDicken 1997).

Skole et al. (1997) have proposed an international system for monitoring land cover change which includes studies in specific locations for field validation and accuracy assessments for the large area analyses; these sites could also be useful for evaluating project impacts, if integrated with the approach described next.

Low-level remote sensing. Using aerial photography, videography, and orthophotographs, photographs of land areas can be taken on an annual basis to see whether the project is proceeding according to design.[2] Field/site measurements and ground truthing will also need to be conducted periodically.

4.2.4. Field/site measurements

Field/site measurements include two types of techniques (biomass surveys and destructive sampling) which can be used together in monitoring carbon in forestry projects (Box 4).

[1] The Face Foundation was set up by Sep (the Dutch Electricity Generating Board) to fund projects to sequester some of the carbon dioxide emitted into the atmosphere by the burning of fossil fuels when generating electricity in the Netherlands. Face stands for Forests Absorbing Carbon dioxide Emissions.

[2] An orthophotograph is a vertical aerial photograph from which the distortions due to varying elevation, tilts and surface topography have been removed, so that it represents every object as if viewed directly from above, as in a map.

Box 4

Field/Site Measurement Example

The Reduced-Impact Logging (RIL) Project, a pilot carbon offset project in Sabah, Malaysia, was initiated in 1992 when a power company provided funds to a timber concessionaire to implement timber-harvesting guidelines in a commercial forest reserve. The rationale for the offset is that when logging damage is reduced, more carbon is retained in living trees and, because soil damage is minimized, forest productivity remains high. It is estimated that logging damage to the remaining biomass can be reduced by as much as 50% through pre-cutting vines, directional felling, and planned extraction of timber on properly constructed and utilized skid trails. Other benefits include the preservation of biodiversity and reduced susceptibility to weed infestations and destructive fires.

Evaluation method: To estimate the carbon benefit associated with implementation of harvesting guidelines, a monitoring program was developed based on computer modeling and simulation, as well as field studies for measuring carbon stocks and flows. Prior to logging, four logging units (30-50 ha each) were randomly selected from the 450 ha pilot project area; four additional logging units were randomly selected from an adjacent area to be logged conventionally Within each unit, 20-40 permanent plots (1600 m^2) were established for pre- and post-harvest measurements. Trees within the plots were tagged, mapped, measured (diameter at breast height, dbh) and identified to species or timber species group. Above-ground tree biomass was estimated allometrically using tree inventory data and stem volume-dbh relations and a biomass expansion factor. Below-ground biomass was measured using pits for coarse roots and cores for fine roots. After logging, permanent plots were revisited, and tagged trees were classified by type and degree of damage. From the damage assessment data, the following parameters were estimated: timber volume extracted; necromass produced from harvested trees; necromass produced from trees destroyed during harvesting; and necromass produced from damaged trees that died within the first 8-12 months after logging. Soil disturbance was mapped and measured in the eight logging units that contained permanent plots. Trees in permanent plots were re-measured three years after logging and are scheduled to be re-measured every five years.

Evaluation concerns: the models chosen for calculating biomass were expected to provide reasonable predictions for trees up to 300 cm dbh, but few data were available for large diameter trees: additional biomass data for large trees from tropical wet and moist forests are needed to improve biomass estimates for old growth forests. For the purposes of monitoring carbon offset projects in natural forest, direct sampling of coarse roots, unless conducted at a relatively high intensity, may not provide a biomass estimate with the desired level of precision. In this study, coarse roots contributed disproportionately to the variance in the estimate for pre-harvest biomass and, consequently, to the difference between the two methods in necromass produced.

Findings: Prior to logging, total plant biomass was about 400 Mg ha^{-1}; root biomass represented 17% of the above-ground biomass. During the first year after logging, the mean difference between RIL and conventional logging areas in necromass produced per ha was 86 Mg; about 62% of the difference was due to more trees killed in conventional as compared to RIL areas. Fifty-nine percent of the total biomass was in trees (≥60 cm dbh), placing particular importance on the reliability in estimates of variables related to big trees. The use of a simple factor adjustment to convert above-ground biomass to total biomass may be a reasonable approach to estimating carbon benefits for offset projects when resources for monitoring are limited and below-ground biomass is unlikely to be a major contributor to the carbon benefit.

Sources: (1) Pinard, M. and F. Putz. 1997. "Monitoring Carbon Sequestration Benefits Associated With a Reduced-Impact Logging Project in Malaysia," *Mitigation and Adaptation Strategies for Global Change*, 2(2-3): 203-215; (2) Pinard, M., F. Putz, J. Tay and T. Sullivan. 1995. "Creating Timber Harvesting Guidelines for a Reduced-Impact Logging Project in Malaysia," *Journal of Forestry* 93:41-45; (3) Pinard, M. and F. Putz. 1996. "Retaining Forest Biomass by Reducing Logging Damage," *Biotropica* 28:278-295; (4) Jepma, C. 1997. "Reduced-impact Logging in Indonesia," *Joint Implementation Quarterly* 3(3): 2.

Biomass surveys

Biomass surveys can include one or more of the following methods: research studies; surveys; the monitoring of wood production and end products; and forest inventories. Research studies use intensive data collection and analysis methodologies to typically test research hypotheses. Surveys of project field activities are conducted to see what was actually implemented in the project. This type of monitoring would provide useful data for the evaluation of GHG reduction and sequestration projects, especially if the surveys are combined with other approaches. The monitoring of wood production and end product data is needed to develop historical and trend data for the development of accurate baselines. An account needs to be made of what happens to the wood once it is felled or trees and branches die. If dead wood is regularly collected, it should be measured and its use recorded.

Carbon inventories can be performed at virtually any level of precision desired by inventory sponsors and provide flexibility in the selection of methods, depending on the costs and benefits of monitoring. Monitoring systems need to assess the net difference in each carbon pool for project and nonproject (or pre-project) areas over a period of time. By comparing these changes in the project area to changes in pools unaffected by project activities (i.e. comparison plots), the monitoring effort can assess the impact of the project on carbon storage. Detailed biomass measurement methods can be found in MacDicken (1998).

Above-ground woody biomass. Two approaches are commonly used for assessing the total above-ground biomass of forests (defined as biomass density when expressed as dry weight per unit area): (1) the first approach is based on the use of existing measured volume estimates (volume of biomass per hectare) converted to biomass density (tons/hectare) using a variety of tools; and (2) the second approach directly estimates biomass density using biomass regression equations that relate oven-dry biomass per tree as a function of a single or a combination of tree dimensions (Brown 1997). The regression equations are applied to stand tables or measurements of individual trees. The advantage of this second method is that it produces biomass estimates without having to make volume estimates, followed by application of expansion factors to account for non-inventoried tree components. The disadvantage is that only a few inventories contain stand tables for small diameter classes for all species. The UN Food and Agricultural Organization has recently published a primer on using these two approaches, including a discussion of the limitations of the approaches (Brown 1997).

Below-ground woody biomass. Roots store carbon and contribute to the build-up of organic soil carbon. The amount of below-ground biomass can be significant: e.g., the ratio of roots to above-ground biomass (i.e., the root:shoot ratio, or R/S) is approximately 25% (Cairns et al. 1997), and others have estimated that approximately one-third of the mass of a tree is below ground (World Bank 1994a). Hence, it may be necessary to measure tree roots — either on the plots or on trees felled outside the project area— to obtain ratios between above- and below-ground woody biomass.

Calculating carbon storage in woody biomass. Once stand, total tree volume, or weight has been estimated, this measure must be converted into organic carbon weight. There is very little variation in chemical composition of all wood species and on an ash free, moisture free (bone dry) basis, approximately 50% of wood by weight is carbon, 6% is hydrogen, and 44% is oxygen (World Bank 1994a). Although the chemical composition of wood does not vary much, density and moisture content vary considerably by species (e.g., coniferous wood species are generally much less dense than hardwood species). Density can be determined by taking pieces of wood of known dimensions, weighing them, subtracting the weight of water, and dividing the volume into the bone dry weight. Moisture content can be measured by weighing the wood as received and reweighing it after it has been dried in an oven until its weight is constant. Alternatively, a moisture content meter can be used which will give a direct reading of moisture content.

Soil carbon. There is no official internationally agreed upon method for monitoring changes in soil carbon. For most lands, soil is usually a greater store of carbon than is biomass tissue, with the most carbon found in forest soils, followed by grassland soils and arable agricultural soils (Bouwmann 1990; World Bank 1994a). Soil accumulation is a function of soil bulk density, which is a function of other parameters, such as the rates of deposition, decomposition, and translocation[1]. Carbon may be lost from some soils under some forest management schemes: e.g., agroforestry projects will disturb the soil, speeding up heterotrophic decomposition which is the main route for carbon to return to the atmosphere from organic matter, and forest management projects on erosion prone areas will lead to reduced soil carbon through translocation.

The buildup of organic carbon in the soil needs to be measured throughout the project site, down to a depth of 30 cm, since land use change has the greatest effect on the upper soil layers (IPCC 1996;

[1] As a consequence of root growth and subsequent decomposition, litter fall and decomposition, microbial degradation and synthesis, mixing by soil fauna, and moisture and temperature cycles, soil organic carbon is allocated over time to different "pools" that are variously defined on the basis of relative recalcitrance which, in turn, governs residence and turnover times. For one typology of soil carbon pools, see Eswaran et al. (1995).

MacDicken 1997).[1] Ideally, soil samples should be taken at permanent sample sites in different age and land use classes, and the buildup of soil carbon recorded. The carbon content of the soil can be calculated using a Leco furnace (which measures carbon by high temperature ignition in a stream of oxygen), a thermal conductivity detector (for separating carbon and nitrogen), or chemical treatment (e.g., Walkley-Black method) (Allison 1975). The potentially high cost of measuring soil carbon may suggest that consideration of changes in soil carbon in many forestry projects is not economically prudent.

Forest products. The long-term effectiveness of wood products as a stock for carbon depends on the uses of the wood produced through project activities. The more durable the wood product, the greater the project's carbon storage effect in the medium and long term. However, carbon stored in wood is obviously not stored permanently; organic compounds usually decay and some will ultimately reappear as GHG emissions. A monitoring and evaluation system to measure post-harvest carbon storage, particularly for medium to highly durable products, could allow reporting of additional carbon and improve the economics of projects that seek to grow higher value timber (Brown et al. 1998; MacDicken 1997 and 1998; Winjum et al. 1998).

Although forest products are not accounted for in the International Panel on Climate Change's 1996 Revised Guidelines (see Section 2.4), an account should be made of what happens to the wood once it is felled or trees and branches die.[2] If dead wood is regularly collected, it should be measured and its use recorded. If it is used as firewood, it may result in lower GHG impacts than if it is left to decompose (due to methane emissions from decomposition). When logs, pulpwood, cord wood and chips are taken to a factory, a record should be made of the fate of this wood: e.g., waste, pulp and board products, animal bedding, fuel within the factory, fuel by households, industry, etc. Similarly, the kinds and quantity of finished products should be recorded: e.g., furniture, recycled paper, or substitute for fossil fuel.

Given the inherent difficulty in determining the exact fate of wood products after they leave the forest or project area, another approach is to determine the proportion of timber that is converted into different products, and use general default values to estimate their average lifetime and decay rates (EcoSecurities 1998). For example, in an analysis of the carbon costs and benefits of silvicultural plantations in Brazil, all pulpwood was assumed to go into the short-term wood

---

[1] Deeper soil layers can also have appreciable carbon stocks, particularly in tropical soils, but they are generally much less impacted by changes in land use/management than are topsoil layers (IPCC 1996).

[2] An IPCC expert meeting in Dakar, Senegal, examined a range of approaches for estimating the emissions and removals of CO_2 from forest harvesting and wood products (IPCC 1998).

products pool (average residence time of 0.5 years), while sawlog wood was assumed to be allocated with 40% entering the short-term pool, 50% the medium-term pool (average residence time of 5 years), 8% the long-term pool (average residence time of 50 years), and 2% the very long-term pool (average residence time of 500 years) (Fearnside 1995).

Destructive sampling

Destructive sampling is the oldest methodology for estimating biomass density at a site. It involves selection of representative sites in the ecosystem (usually a few square meters each, and in a few rare cases as large as one hectare each). All the vegetation is uprooted and the pertinent parameters obtained, e.g., volume, weight at different moisture contents, proportions of various components like branches, stem and roots, and chemical composition of the biomass. Detritus is also collected and similarly analyzed. This is usually accompanied with similar measurements of parameters of interest in the soil profile, including soil layers, structure, texture and cation exchange capacity, organic carbon, inorganic nutrients, etc.

4.2.5. Sampling

Sampling allows overall project performance to be assessed based on the performance of a manageable number of plots. For large, heterogeneous areas, a multi-stage approach may be appropriate, in which each stratum is divided into primary sampling units which are then subsequently divided into secondary sampling units. The type and intensity of sampling depends on the variations within each stratum. Biomass sampling studies typically aim for estimates of biomass weight or volume accurately to within ± 15% with a relatively high confidence (e.g., 90 or 95%) (World Bank 1994a); biomass estimates within 2-10% of the true value are also realistic (MacDicken 1997).

A universally accepted level of precision for estimates of carbon benefits does not currently exist. As a general rule, the cost of a monitoring program is related to the precision of the estimate of the carbon benefit: the higher the precision, the higher the cost of measurement. To a certain extent, the market value of carbon sequestered in carbon offset projects will determine the level of precision that is cost-effective. Some experts suggest that a reasonable target for the precision of a project's carbon benefit is a standard error of 20-30% of the mean (EcoSecurities 1998). Another option would be to adjust the carbon claims by discounting the standard error of measurements. Finally, it is unlikely that a common level of precision will be used for each of the significant carbon pools and flows.

The use of permanent sample plots is generally regarded as a statistically superior means of evaluating changes in forest conditions (MacDicken 1998). Permanent plots allow reliable and efficient assessment of changes in carbon fixation over time, provided that the plots represent the larger area for which the estimates are intended. This means that the sample plots must be subject to the same management as the rest of the project area. The use of permanent plots also allows the inventory to continue reliably over more than one rotation. Finally, permanent plots permit efficient verification at relatively low cost, compared to those that use temporary plots or plotless methods: a verifying organization can find and measure permanent plots at random to verify the design and implementation of a project's carbon monitoring plan. The size of the permanent plots will depend on the heterogeneity of the site.

Instead of conducting a census, three sampling approaches may be used: simple random sampling, systematic sampling, and stratified random sampling. For carbon inventory, stratified random sampling is generally preferred, since this often yields more precise estimates for a fixed cost than the other options (see Box 3) (MacDicken 1998). Stratified random sampling requires dividing the population into nonoverlapping groups. Each stratum can be defined by vegetation type, soil type, or other parameters for sampling purposes.

Useful tools for defining strata include satellite images, aerial photographs, and maps of vegetation, soils or topography (see Box 3). These should be combined with ground measurements for verifying remotely-sensed images. A geographic information system can be used to determine stratum size and the size of exclusions or buffer zones.

MacDicken (1998) provides a spreadsheet for inventory decisions which calculates sample sizes using standard formulas based on measured variation for the carbon pool to be sampled. Two approaches are proposed: (1) sample plot allocation based on fixed precision levels; and (2) optimum allocation of plots among strata given fixed inventory costs.

4.2.6. Application of forestry monitoring techniques

The unique features and diversity of forestry projects, the monitoring domain and socioeconomic issues pertaining to forestry projects, and the variety of carbon pools that might be impacted by forestry projects makes the monitoring and evaluation of forestry projects very challenging. While forestry projects offer the potential for significant carbon sequestration, the verification of carbon credit claims will necessitate significant technical and financial resources. A variety of monitoring techniques are available for forestry projects (i.e., modeling, remote sensing, and field/site measurement) for determining the amount of carbon sequestered by forestry projects, each having its own advantages and disadvantages (Table 2).[1] One of the key decisions that will need to be made will be determining the optimal level of costs for implementing these techniques.

We expect the use of these techniques will vary by the size of the project area, region, type of forest, and the purpose of the project (e.g., to protect forests, supply energy, or provide wood products). Using some of these criteria, we provide a table classifying monitoring techniques by forestry typology (Table 3). The threshold for distinguishing small from large projects is not known and will be left to the project developer to decide.

[1] As noted in Section 2.3, the monitoring and evaluation of the impacts of biomass energy projects will rely on the methods described in this section as well as methods used in monitoring and evaluating energy-efficiency projects (Vine and Sathaye 1999).

Table 2. Advantages and Disadvantages of Forestry Monitoring Techniques

Techniques	Advantages	Disadvantages
Modeling	Relatively quick and inexpensive. Useful for baseline development. Can be used for bioenergy projects. Most useful as a complement to other methods.	Relies on highly simplified assumptions. Need to be calibrated with onsite data.
High-Level Remote Sensing	Provides relatively rapid regional-scale assessments of land cover, land use, and green vegetation biomass. Useful for monitoring leakage.	Time and knowledge needed to transform spectral classifications into accurate land use or land-cover classifications. Access to high-quality imagery may not be available during certain seasons or due to sun angles. Has not been used to measure carbon. Can be quite expensive.
Low-Level Remote Sensing	Complements high-level remote sensing. Useful for monitoring leakage.	In test phase. Less expensive than high-level remote sensing.
Field/Site Measurements	Useful for determining what was actually implemented in project and for tracking fate of wood products. Flexible in selection of methods and precision. Peer reviewed and field tested systems available. Using control plots, can calculate net carbon sequestration.	May be more expensive than other methods.

Table 3. Forestry Monitoring Methods by Forestry Project Type

(✔ = applicable; blank = not applicable)

	Carbon conservation		**Carbon sequestration and storage**		**Carbon substitution**
Methods	**Small Project**	**Large Project**	**Small Project**	**Large Project**	
Modeling	✔	✔	✔	✔	
Remote Sensing		✔		✔	✔
Field/Site Measurements	✔	✔	✔	✔	✔

The monitoring process is an evolving process that is expected to change over time. For example, in the early stages of a reforestation project, monitoring will most likely be visual, with widely scattered sample plots, because the focus is on seedling survival, not on carbon sequestration (carbon sequestration is likely to be minimal for the first five years of the project). Soil measurements would also not be necessary in the first few years of the project. Once the stand is established, monitoring would switch to growth evaluation and soil monitoring. The results from the monitoring would provide feedback to reporting and financial planning models (i.e., calibration and real-time adjustment).

4.2.7. Quality assurance guidelines

Implementing data collection and analysis methods is both an art and a science, and there are known problems associated with these methods. Thus, simply adhering to minimal standards contained in these guidelines is no guarantee that an evaluator is doing a professional job. Accordingly, we have included Quality Assurance Guidelines (QAG) that require evaluators and verifiers to indicate specifically how basic methodological issues and potentially difficult issues were addressed (see Appendices B and C).[1] The guidelines are contained in two tables, covering all of the data collection and analysis methods.

The QAG should be seen as practice and reporting standards, rather than highly prescriptive methodological standards: the QAG require evaluators to describe how certain key issues were addressed rather than to require them to address these issues in a specific way. Adherence to such guidelines still allows the methods to be shaped by the interaction of the situation, the data, and the evaluator.

The QAG are to be used in three ways. First, they are included in the Monitoring and Evaluation Reporting Form (Appendix B), so that evaluators will know that they will be held accountable for conducting a sound analysis. Second, they are included in the Verification Reporting Form (Appendix C), so that policymakers and other stakeholders could review a verification report and quickly assess whether the evaluator addressed the most basic methodological issues. This is especially important since most stakeholders do not have the time nor personnel to carefully scrutinize every written evaluation report, let alone attempt to replicate the results of all of these studies. The details of how evaluators addressed these methodological issues should be contained in the very detailed documentation that would be in the technical appendix of any evaluation report, or in working papers. Finally, the QAG can be used to create a common language to facilitate communication among project developers, evaluators, verifiers, policymakers, and other stakeholders.

Evaluators and verifiers should consider the issues involved in conducting these methods, some of which have been described previously, and which are listed in Table 4 and described in more detail in Appendices B and C. The column headings refer to the data collection and analysis methods described in Section 4.2. The rows refer to the types of issues to be considered when addressing each method. Examples of each of these issues are mentioned below:

- **Calibration**: e.g., were the assumptions and estimated results of models compared and adjusted to actual data?
- **Sample and sampling**: e.g., what kind of sampling design was used?
- **Data type and sources**: e.g., what was the source of the data and the methods used in collecting data?

[1] These guidelines are primarily based on the QAG that were developed for the California Demand-Side Management Advisory Committee (CADMAC) (Ridge et al. 1997). In theory, the QAG could be used in the estimation stage, but are not included in the Estimation Reporting Form.

- **Specification and error**: e.g., what kind of errors were encountered in measuring variables and how were these errors minimized?
- **Outliers**: e.g., how were outliers and influential observations identified and handled?
- **Missing data**: e.g., how were missing data handled?
- **Weather**: e.g., what was the source of weather data used for the analysis?
- **Comparison group**: e.g., how was a comparison group defined for estimating net carbon sequestered?
- **Measurement duration**: e.g., what was the duration and interval of measuring carbon?
- **Variance**: e.g., how were confidence intervals derived?

Table 4. Quality Assurance Issues for Data Collection and Analysis Methods[1]

(✔ = applicable; blank = not applicable)

	Modeling	Remote Sensing	Field/Site Measurements
Calibration	✔		
Sample and sampling		✔	✔
Data type and sources	✔	✔	✔
Specification and error		✔	✔
Outliers		✔	✔
Missing data	✔	✔	✔
Weather	✔	✔	✔
Comparison group		✔	✔
Measurement duration		✔	✔
Variance	✔	✔	✔

1 Quality assurance issues (rows) are described in Appendices B and C, and the data collection and analysis methods are described in Section 4.2.

4.2.8. Project leakage and positive project spillover

In the beginning stages of a project, project leakage and positive project spillover are likely to be modest, so that the MERVC of such impacts may not be a priority. These effects are also likely to be insignificant or small for small projects and for certain types of projects. Under these circumstances, it may be justified to disregard these impacts. This would help reduce MERVC costs. As the projects become larger or are more targeted to market transformation (see Section 4.2.9), these impacts should be evaluated. As an example, in the Rio Bravo Carbon Sequestration Pilot Project, secondary impacts were deemed to be significant if the impacts resulted in an alteration in emissions of 5,000 tC/yr or above (i.e., 20% of the 1 million tC estimated to be sequestered through the purchase of forested land, or 200,000 tC, divided by the 40 years of the project life) (Programme for Belize 1997). Furthermore, to be "clearly and directly" attributable to the project, the secondary impacts had to manifest themselves within 1 year (Programme for Belize 1997); for the evaluation of forestry projects, longer periods (e.g., 5 years) may be necessary.

4.2.9. Market transformation

The focus of most evaluations of market transformation projects is on market effects (Eto et al. 1996; Schlegel et al. 1997): e.g., the effects of forestry projects on the structure of the market or the behavior of market actors that lead to increases in the adoption of forestry products, services, and/or practices. In order to claim that a market has been transformed, project developers and evaluators need to demonstrate the following (adapted from Schlegel et al. 1997):

- There has been a change in the market that resulted in increases in the adoption and penetration of forestry technologies and/or practices.
- That this change was due at least partially to a project (or program or initiative), based both on data and a logical explanation of the program's strategic intervention and influence.
- That this change is lasting, or at least that it will last after the project is scaled back or discontinued.

The first two conditions are needed to demonstrate market effects, while all three are needed to demonstrate market transformation. The third condition is related to the discussion on permanence: if the changes are not lasting (i.e., they do not persist), then market transformation has not occurred. Because fundamental changes in the structure and functioning of markets may occur only slowly, evaluators should focus their efforts on the first two conditions, rather than waiting to prove that the effects will last.

To implement an evaluation system focused on market effects, one needs to carefully describe the scope of the market, the indicators of success, the intended indices of market effects and reductions in market barriers, and the methods used to evaluate market effects and reductions in market barriers (Schlegel et al. 1997). Evaluation activities will include one or more of the following: (1) measuring the market baseline; (2) tracking attitudes and values; (3) tracking sales; (4) modeling of market processes; and (5) assessing the persistence of market changes (Prahl and Schlegel 1993). As one can see, these evaluation activities will rely on a large and diverse group of data collection and analysis methods, such as: (1) surveys of customers, forestry companies, furniture manufacturers, government organizations, etc.; (2) analytical and econometric studies of cost data and sales data; and (3) process evaluations.

4.3. Re-estimating the Baseline

During project implementation, the baseline needs to be re-estimated, based on monitoring and evaluation data collected during this period. In some cases, allometric equations for estimating carbon emissions may be used, but only under special conditions (see Section 3.2). In the re-estimation of a baseline, free ridership needs to be examined.

4.3.1. Free riders

The most common method of developing an estimate of free riders is to ask project developers what they would have done in the absence of the project (also referred to as "but for the project" discussions). Based on answers to carefully designed survey questions, project developers are classified as free riders (yes or no). For example, would the construction of an energy-efficient sawmill have been constructed without a joint implementation project. There are at least two problems in using this approach: (1) very inaccurate levels of free ridership may be estimated, due to questionnaire wording;[1] and (2) there is no estimate of the level of inaccuracy, for adjusting confidence levels. Nevertheless, some interviewing of project developers needs to be conducted for deriving estimates of free ridership.

[1] For example, in an analysis of free ridership in a high-efficiency refrigerator program, estimates of free ridership varied from 37% to 89%, depending on questionnaire wording (Boutwell et al. 1992).

4.3.2. Comparison plots

For some projects, the comparison of the amount of carbon storage achieved under a project with the amount that would have been achieved without the project requires monitoring the project area as well as nonproject comparison sites prior to project startup. One can have comparison plots within the project area or outside the project area to supplement the sites within the project area. To establish the internal validity of the evaluation results, the comparison plots must be similar enough to the project area so that they can serve as a proxy for the project area under the assumption that the project was not implemented.[1] Similarity can be established on the basis of the key factors that determine biomass productivity: rainfall, temperature, insolation, soil characteristics, species and land management. Land management is the most difficult criterion to meet since it could diverge significantly between comparison site and project areas. By selecting comparison plots within the project area, these divergences can be eliminated or minimized. Also, there is no general way to ensure that the comparison plots will remain valid throughout the life of the project; special care and monitoring are needed.

[1] This is particularly important when trying to estimate deforestation rates for protected areas. The estimation of deforestation rates is critical in establishing project baselines, and slight changes in the estimates of deforestation can significantly affect the amount of carbon saved by a carbon offset project (see Busch et al. 1999).