

Evaluating the Cost-Effectiveness of Collection Management: A Methodological Framework

(Draft - Not for citation)

Daniela Horna and Melinda Smale

with contributions from:

Bonwoo Koo, Ehsan Dulloo, Thomas Payne, Ruaraidh Sackville-Hamilton, Jean Hanson, Ines Sanchez, Daniel Debouck, Ines van der Houwe, Hari Upadhyaya, Alexandra Jorge

1. Introduction

Conserving germplasm is a long-term activity that requires a long-term perspective, but genebanks are typically funded on a short-term basis (Koo et al. 2004). Phase 1 of the project “Collective Action for the Rehabilitation of Global Public Goods in the CGIAR Genetic Resources System” (GPG), the thorough analyses of conservation costs previously undertaken by the Systemwide Genetic Resources Program (SGRP) (summarized in Koo et al. 2004), and other economics research about the benefits of genebanks have demonstrated the importance of sustained funding and the high expected benefits of ex-situ conservation relative to costs, assuming “good practices” (summarized in Smale and Drucker, 2007; Smale and Koo 2003; for the case of a large national genebank, see also Day-Rubenstein et al. 2006).

The expansion of genebank collections from the 1970s through the 1990s led to management challenges. These included the duplication of accessions, backlogs in regeneration, and insufficient or untimely provision of information to users (Altoveros and Rao 1998; Engels and Rao 1998; Koo and Wright 2008)¹. At the same time, there

¹ Fowler and Hodgkin (2004) report that between 1974 and 1996, the number of long-term storage facilities in the world grew from five or six to 76, with an estimated 6.2 million accession housed by gene banks located in 137 countries. Experts estimated that by the mid-1990s, only five percent of the rice, maize, and wheat gene pools remained unrepresented among these accessions. These authors caution that : a) coverage is much lower for many crops b) it is not possible to catalog a crop’s gene pool with any precision and c) while some duplication is necessary to safeguard accessions, the redundancy of materials could be

was increasing recognition that integration and coordination of the collections as a global system offered important functional and economic advantages. In 1995, SGRP commissioned an external review of the CGIAR genebanks to provide an assessment of what was needed to meet conservation standards. The first phase of the GPG project addressed the main recommendations from this review, including amelioration of genebank facilities and genebank procedures.

The second phase of the project (GPG2) builds on the progress made in the first phase, with a focus on establishing good standards and practices in genebank operations and encouraging a systems perspective. The current challenge, as viewed by those engaged in this project, is not to increase the numbers of accessions, but to ensure the quality, security, accessibility and sustainability of the in-trust collections. An underlying assumption is that a better allocation of resources will lead to better performance.

Presently, genebanks operate according to protocol manuals. These manuals are recognized as useful, but they do not ensure best practices. Clearly, establishing a set of best practices based on performance indicators will increase the probabilities of delivering a high quality product because the chances of making a technical mistake will decrease. What is less clear is how these performance indicators relate to genebank expenditures. As genebank managers point out, in any case, there is a need to examine the cost-effectiveness of operations (output per cost).

The goal of the economics task in the GPG2 project is to develop and disseminate a computerized tool that will support strategic decision-making by genebank managers. The objectives of this document are to a) provide a conceptual framework for the tool and b) demonstrate how the tool can be used to evaluate the effects of decisions on the allocation of resources across operations. Effects of decisions are illustrated by two types of outputs: a) cost summary reports and b) sensitivity analysis with simulations. Thus, it is expected that genebank managers will be able to apply the tool to answer management questions and craft strategies in pursuit of good practices or to enhance their

substantial. Regeneration of large collections is costly. Thus, short-term budgetary constraints could endanger the longer-term viability of such collections.

performance. Eventually, the tool could be generalized in order to explore the effects of resource allocation decisions within an integrated genebank system.

This study is divided into six sections. The second section discusses some fundamental concepts on which the framework is based. Section 3 outlines a decision support tool that builds on the work previously undertaken by the SGRP. This section describes the type of information genebank managers need to apply to the tool and the outputs that can be produced. Notice that a single set of cost data (representing one point in time) allows us to minimize cost only with respect to the technology and set of practices represented by those data. To draw conclusions concerning optimal allocation of resources within a single bank over time, and among banks, additional points corresponding to other technologies and practices are needed. Section 4 and 5 presents two additional analyses that can be conducted with additional observations: sensitivity analysis and regression analysis. In the last section, we discuss some considerations for the implementation of this framework as well as next steps for this activity.

2. Defining concepts

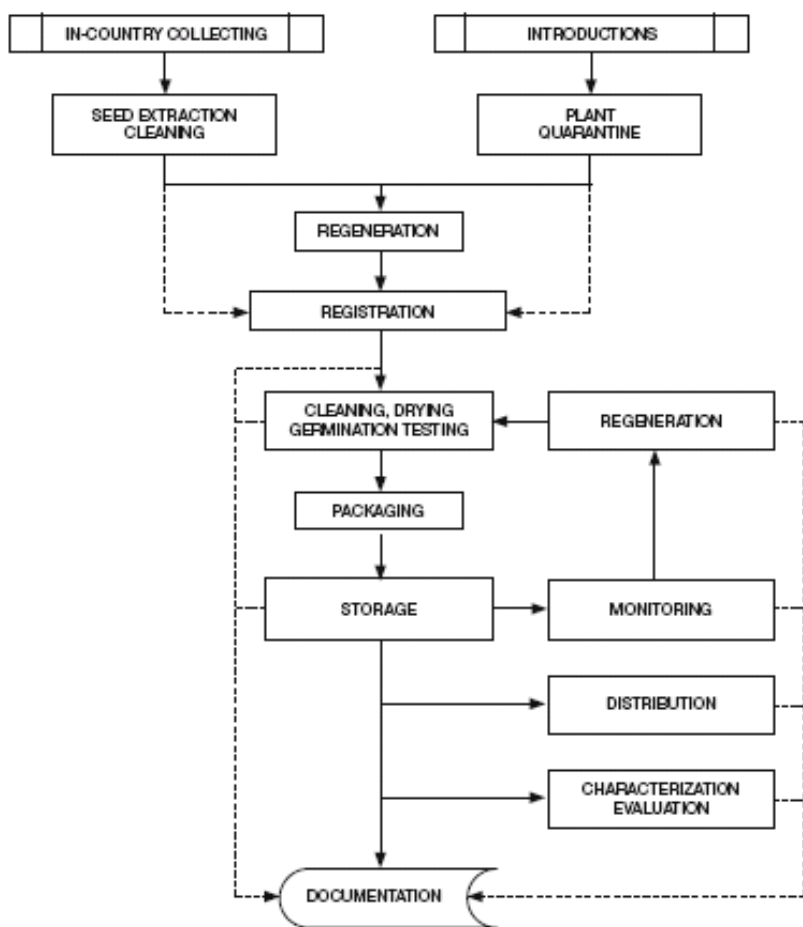
a) Genebank Operations

There is no disagreement over the main purpose of a genebank, which is to conserve genetic material and make it available to users. However, a review of genebank protocols suggests that agreement has not yet been reached on a general classification of activities and related terminology (Rao et al. 2006; Taba et al. 2004).

Pardey et al. (2001) and Koo et al. (2003) group genebank activities into operations performed to reach genebank objectives. Orienting their description toward “best practices,” Calles et al. (2007) classify genebank activities and inputs according to the specific objectives. Many operations are comparable across Centers, but other activities are specific to reproduction system of the crop, such as propagation and multiplication strategies. Seed propagated crops like wheat or rice are the easiest to handle and can be conserved for longer periods than clonal crops like cassava or banana (Rao et al. 2006). Figure 1A depicts some of the main activities of a seed germplasm bank. Figure 1B presents a generic flow of operations.

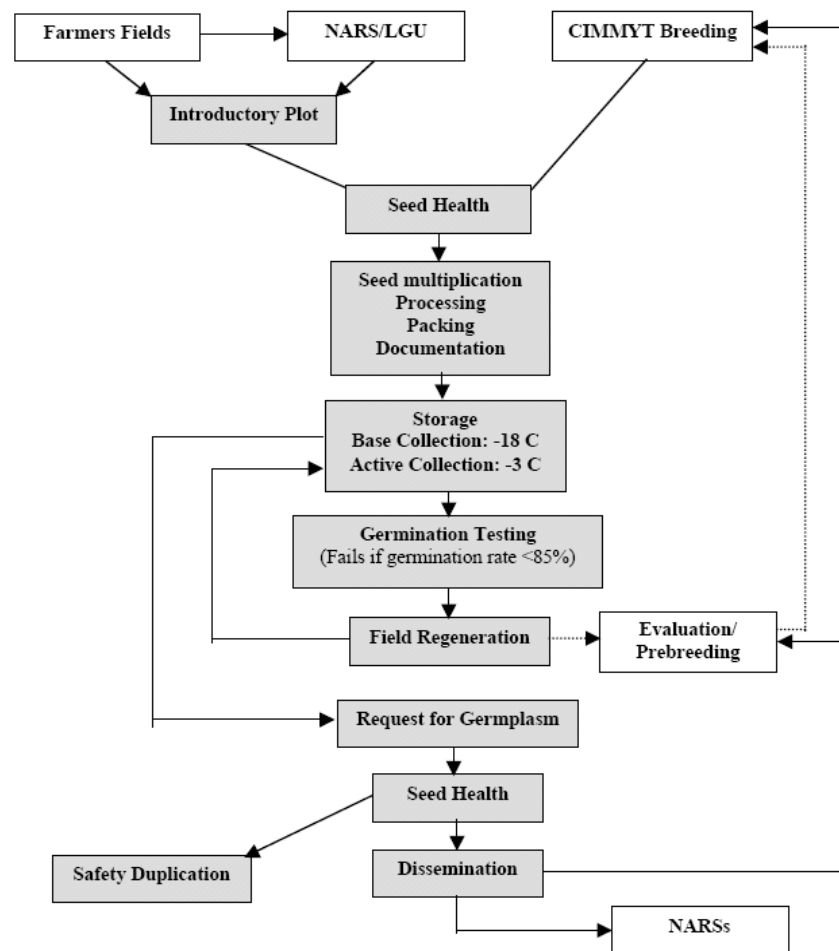
Figure 1. Genebank Flow of Operations

A) General sequence of operations in a seed genebank



Source: Rao, N.K. et al. 2006

B) Generic flow of operations in a CG genebank



Source: Pardey et al. 2001

In this study, an operation is understood as a cluster of activities and a number of operations are performed in order to reach a genebank objectives and thus genebank goals. Conservation and use of the genetic material are the two main goals of a germplasm bank. Specific objectives for achieving better conservation of genetic materials are:

- To cover the gene pool as much as possible
- To ensure the security (physical security and viability) of the genetic material
- To maintain its genetic integrity

Specific objectives for achieving a wider use of genetic materials are:

- To ensure the availability of the material to users
- To distribute the material
- To provide information

b) Best Practices

The issue of quality standards is central to the management of any genebank. Genebanks in the CG system have operated with two sets of conservation standards: acceptable and preferred. Acceptable standards are considered to be minimal but adequate, while preferred standards guarantee better and safer conservation conditions. Evidently, meeting the preferred standards is more costly. Acceptable standards have been more frequently adopted as a consequence of budget constraints, leading to wide variation in quality standards across centers.

The CG genebanks are now directed toward “best practices,” which is a more dynamic, less easily defined concept of quality management. Genebank managers have not yet reached a consensus regarding the operational meaning of best practices. In some cases, “best practices” are viewed as activities that mitigate the risks that impede the achievement of objectives (conservation and use). In other cases, “best practices” are simply understood as the most effective practices given the technology that is currently available to the research center. In this study, we understand best practices as the costs incurred in order to reduce the chances of mistakes in technical procedures or in the delivery and distribution of genetic materials and related information to use. Different managers may have different degrees of tolerance of error probabilities, although they might agree on a standard threshold. Therefore the concept of best practices is directly

linked to that of risk management. Ideally, implementing best practices based on performance levels as expressed by a set of indicators should minimize spending subject to an acceptable threshold of risk, a current conservation technology, and a current organization of collections. This would be one point of optimal resource allocation. Other optima would correspond to other conservation technologies, other risk thresholds, or a different organization of collections in the genebank system.

c) Performance indicators

Performance indicators measure the quality of an operation or a system in quantitative terms. Good performance indicators should be simple and measurable, while capturing the essential features of a complex system. Identifying performance indicators for the CG genebanks is another activity (No. 6.1.2) of the GPG project². The task group has identified three kinds of performance indicators: technical operations indicators, impact and relevance indicators, and international collaboration indicators. For this study, only technical and impact indicators are considered to be relevant.

The framework proposed here would allow managers to discern how they might improve performance through re-allocating resources, or how they might maintain performance despite budget constraints. However, performance indicators must be defined in order to continue implementing the framework outlined in this document. The initial challenge is to make the right assumptions about the links among performance indicators, input use and costs. These links may not be as intuitive as expected. The variation in life cycle of the different operations conducted in the genebank, the share of resources allocated by genetic material, activity and/or operation diffuse the effect of input use on performance, making it difficult to isolate and establish causal relationships.

² Initial version of performance indicators identified for the GPG project are presented in Annex 1

3. Analysis of Genebank Costs³

A review of CGIAR genebanks in the mid-1990s demonstrated the need for upgrading management of the in-trust collections. In response, SGRP organized a series of economic studies to determine the costs of the maintaining collections and proposed upgrade. The GPG project was established to facilitate the upgrade. During the first phase of the project, which started in 2003, in close consultation with several CG genebank managers, Koo et al (2004) compiled and analyzed genebank cost information.⁴

The analytical framework for the cost studies was the micro-economic theory of production (Pardey et al. 2001). A genebank, like a firm, is organized to produce outputs (numbers of accessions characterized, stored, regenerated, etc.). Production decisions involve choosing which outputs to produce in which amounts, with which mix of inputs and input quantities. In the framework of economic decision-making, optimal resource allocation can be achieved either by minimizing the costs of operation given fixed physical resources and existing technology or by maximizing production subject to a fixed budget and existing technology. By duality theory, it has been proven that both approaches produce the same production possibility frontier. The production possibility frontier then traces the points corresponding to efficient resource allocations.

This approach selected by Koo et al. (2004) was cost minimization—for a very important, practical reason. Most of the benefits of genebank collections are public goods whose values are both expensive to estimate and likely to be unreliable estimated (see Smale and Koo 2003). By comparison, the costs of genebank operations are relatively easy to estimate with a fair degree of precision. Pardey et al. (2001) reasoned that if the costs of conserving an accession are shown to be lower than any sensible lower-bound

³ This approach is based on cost minimization approach. In the framework of economic decision-making, optimal resource allocation can be achieved either by minimizing the costs of operation given fixed physical resources and existing technology or by maximizing production subject to a fixed budget and existing technology. By duality theory, it has been proven that both approaches produce the same production possibility frontier. The production possibility frontier then traces the points corresponding to efficient resource allocations.

⁴ These studies provided evidence to enable the Global Crop Diversity Trust to make realistic resource projections for an endowment to support globally important collections of crop diversity in perpetuity, including those held by the Centres (<http://www.sgrp.cgiar.org/CurrentSGRPInitiatives/GPGProject.htm>).

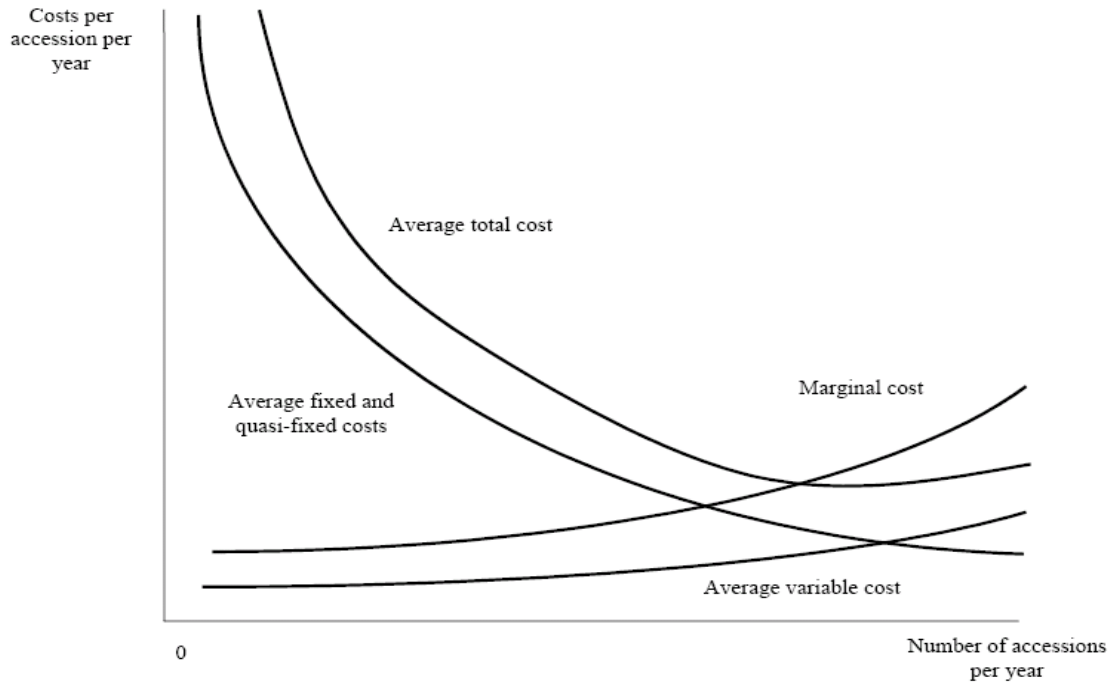
estimate of the corresponding benefits, for many decisions, it may not be necessary to estimate benefits.

The data compiled by genebank managers on input use and expenditures was used to estimate average and marginal cost per unit. Average costs are the costs for the genebank of managing 1 accession. Marginal costs are the increase in total costs from the addition of one more accession to the genebank. Total costs include costs that vary and costs that are fixed in the relevant range of production. Average fixed or quasi-fixed (genebank management) costs normally decline as output increases. A standard assumption of micro-economic theory is that marginal costs initially decline as more is produced in a plant and eventually increase due to diminishing marginal returns to fixed factors (e.g., land, plant). Marginal cost is equal to average total costs when average total cost is at a minimum. Notice however that often genebanks operate below capacity, average costs then represent only upper bounds estimates of the marginal costs. Figure 2 illustrates how average and marginal costs are thought to change with amounts produced (for example, the number of seeds stored, regenerated, disseminated, etc).⁵

The research summarized in Koo et al. (2004) represents only a single year of data for 5 CG Centers. In order to evaluate genebank costs more generally a broader cross-section and longer time series is vital. Unfortunately, genebank operation costs have not been systematically recorded in the CG system. This information must be gathered in a uniform and systematic way in order to facilitate comparison across genetic materials and across centers. The use of the decision support tool, described below, could facilitate the implementation of a system for periodic data collection by managers. Managers can use this information to monitor and evaluate their own performance, and as in input into strategic organizational decisions.

⁵ A few other studies have addressed the issue of cost implication of germplasm conservation, but without an explicit micro-economic framework. For example, Virchow (2003; 1999) used surveys to collect national conservation expenditure for 39 countries and estimated per-accession cost of annual conservation for each country. Burstin et al. (1997) also used surveys, examining the cost associated with sexually and vegetatively propagated species in several French genebanks. The authors calculated the annual and long term costs of each operation. Survey-based studies often suffer of inconsistent responses and excessive aggregation.

Figure 2: Genebank average and marginal cost



Source: Pardey et al. 2001.

Decision Support Tool

A prototype of a **decision support tool** has been developed based on the framework of Koo et al. (2004). The first purpose of the tool is to store detailed input use per operation and generate cost reports. So far, the tool has been developed as an excel file with an introduction sheet, a general information sheet, 3 input sheets (non-labor, labor and capital inputs) and 4 output sheets or reports (see Annex 2). The introductory sheet provides a brief explanation of the purpose of the tool and the framework used to classify activities, inputs and costs. The general information sheet elicits details about the genebank (e.g., genetic material, number of accession managed, etc.) and other factors that affect costs (e.g., discount factor, overhead rate, period for performing operations).

Detailed **input** use and related expenses are entered in the decision tool, dividing the information by type of input (the categories are capital, labor and non-labor). In general, capital inputs are not sensitive to the size of the operation. Capital inputs include infrastructure, such as germplasm storage and genebank facilities, and equipment for

field operations and offices. Variable inputs, on the other hand, are sensitive to size of the operation. Variable inputs include non-labor costs and some labor costs. Non-labor variable costs mainly include inputs consumed on a daily basis, like energy, office and laboratory supplies. Variable labor costs are salaries paid to temporary workers and non-senior staff. Senior scientists and technicians are treated as quasi-fixed inputs. Quasi-fixed inputs are more variable than fixed capital inputs but unlike variable costs, they are not easily apportioned when the size of the operation changes.

All inputs used and expenses must be allocated by operation using rates. For instance, the total energy consumption in a genebank must be distributed among all operations that required energy. Allocation requires expert knowledge about the demands of genebank operations. Genebank managers thus are the persons who, in consultation with their staff, are most able to provide good estimates of allocation rates. Information about inputs is used to determine capital costs, quasi-fixed cost, variable costs, and genebank total costs. Allocation rates disaggregate these costs per operation.

To produce **output reports**, total costs are broken down into capital, variable, and quasi-fixed costs. To illustrate, we used the example of the WARDA genebank in 2006, provided to us by the genebank manager. In addition to a summary overview by crop and input costs (Annex 3), three kinds of output reports can be generated. The first output reports costs per input category, genetic material, and operation, as shown in Table 1. The report provides information about both total costs and average costs per accession. The report also includes a graphic representation of the distribution of total costs (Figure 3). In the current version of the tool, this graph depicts the distribution of costs per input type, but other graphs could be developed based on expressed needs of genebank managers.

The second output corresponds to annual and in-perpetuity average cost per accession classified in terms of either conservation or distribution costs (Table 2). Acquisition, viability testing, duplication, storage, and regeneration are operations that need to be performed in order to conserve an accession. Characterization, storage, regeneration, and dissemination are operations that are necessary in order to be able to distribute an accession. Costs are estimated for both new accessions and existing

accessions, to indicate the additional cost of acquiring new accessions as compared to managing current accessions. The last output summarizes distribution and conservation costs associated with maintaining all existing genebank accessions (Table 3). In this analysis, distribution costs are treated as short-run costs and conservation costs are considered to be “long-run” costs. This report shows the annual and in-perpetuity costs for the genebank. Such information is useful when justifying genebank funding or investment in ex-situ conservation.

While these reports help to understand the structure of genebank costs and their distribution across operations, objectives and over time, nothing can be inferred about the factors that affect these costs. For this reason, although it is possible to compare reports across genebanks, we do not have a picture that enables us to tackle strategic decisions. Two feasible ways to extend the use of the tool are discussed in the following sections.

Table 1: Decision Support Tool Report 1

Activities	Number of accessions	Total capital cost	Total quasi-fixed cost	Total labor costs	Total non-labor costs	Average capital cost	Average quasi-fixed cost	Average variable cost	Total AC
Acquisition	0	8,385.57	463,383.39	174,264.93	0.00	0.00 ⁶	0.00	0.00	0.00
Characterization	204	268,834.82	501,166.10	486,782.39	230.59	1,317.82	2,456.70	2,387.32	6,161.83
Safety duplication	2,500	3,483.56	249,379.21	27,695.76	230.59	1.39	99.75	11.17	112.32
Long term storage	5,000	3,483.56	190,214.09	7,887.35	0.00	0.70	38.04	1.58	40.32
Medium term storage	12,000	11,959.03	507,302.04	10,252.35	99,056.55	1.00	42.28	9.11	52.38
Germination testing	1,105	130,720.53	395,981.45	253,599.05	2,305.94	118.30	358.35	231.59	708.24
Regeneration	1,094	134,960.61	896,158.78	661,111.41	419.40	123.36	819.16	604.69	1,547.21
Seed health testing	100	10,324.99	90,505.09	145,433.75	230.59	103.25	905.05	1,456.64	2,464.94
Dissemination (or distribution)	4,040	10,613.44	731,263.79	359,171.88	461.19	2.63	181.01	89.02	272.65
Information and data management	4,679	47,011.80	1,049,245.45	1,081,217.70	461.19	10.05	224.25	231.18	465.47
Total	N.A.	629,778	5,801,355	3,213,727	123,161	1,678	5,125	5,022	11,825

⁶ In this year there were no new accessions acquired.

Figure 3a:

Cost Distribution for RICE

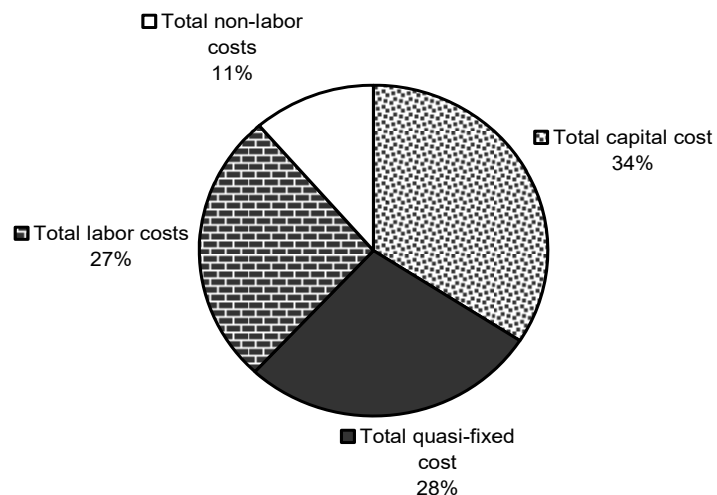
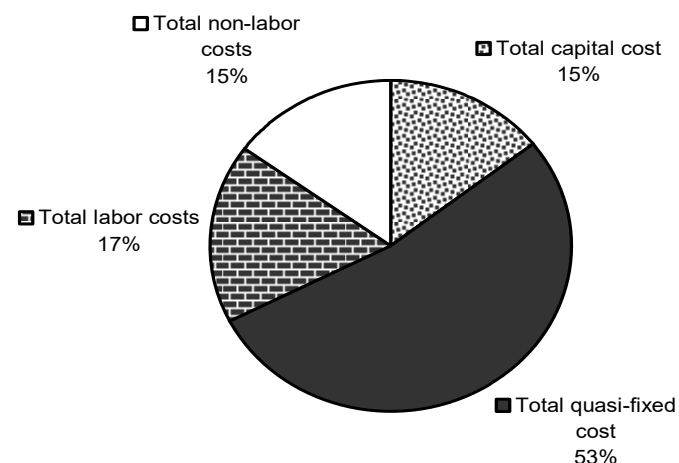


Figure 3b:

Cost Distribution for WILD RICE



These graphs show the distribution of total costs across genetic materials managed in a genebank, in this case rice and wild rice. In the case of rice labor costs (including quasi-fixed costs) account for about half of the costs. For wild rice the rate of labor costs to total cost is much higher. This is probably due to the demand of highly qualified staff for characterization and regeneration of wild rice accessions. Rice accessions also have to be characterized and regenerated but the total number of accession is higher. Total quasi-fixed costs show limited changes in value with operation size. Average quasi-fixed costs decrease with an increased on the number of accessions managed.

Table 2 Decision Support Tool Report 2

Crop: Rice ⁷

Activity	Annual Costs			In Perpetuity	
	Existing accession		New accession (a)	Existing accession (b)	New accession
	No regeneration	Regeneration			
Conservation					
New introduction					
Acquisition			0.00		0.00
Initial viability testing			589.94		589.94
Initial duplication			110.92		110.92
Safety duplication				18.16	18.16
Long-term storage	39.62	39.62	39.62	1,030.13	1,030.13
Viability testing		589.94	589.94	2,722.98	2,722.98
Regeneration (50 yrs.) (c)		711.92	711.92	213.19	1,515.05
Total Cost	39.62	1,341.49	2,042.35	3,984.46	5,987.19
Distribution					
Characterization			4,844.02		2,114.63
Medium-term storage	51.38	51.38	51.38	1,335.99	1,335.99
Regeneration (25 yrs.) (c)		711.92	711.92	1,349.83	2,651.70
Dissemination	270.02	270.02	270.02	7,020.62	7,020.62
Total Cost	321.41	1,033.33	5,877.35	9,706.44	13,122.94

⁷ The information is presented per genetic material kept in the genebank, in this case rice. WARDA also has a wild rice collection.

Table 3: Decision Support Tool Report 3

Crops	No. of acc.	Per-accession cost			Total cost		
		Conservation	Distribution	Total	Conservation	Distribution	Total
Rice	83,600	57	89	146	4,747,043	7,431,347	12,178,390
Noncapital		30	55	84	2,476,961	4,577,309	7,054,269
Capital		27	34	61	2,270,082	2,854,039	5,124,121
Wild Rice	3,200	80	129	209	255,794	412,617	668,411
Noncapital		59	104	163	188,076	334,274	522,350
Capital		21	24	46	67,718	78,343	146,060
All crops	86,800	136.72	217.83	354.55	5,002,836.97	7,843,964.24	12,846,801.21

4. Sensitivity Analysis with Simulations

When it is augmented by sensitivity analysis and simulations, the decision tool can be used to investigate how genebank costs and genebank performance are affected by changes in key parameters. An impediment to analyzing genebank costs across centers is the limited information that is available. One way to overcome this impediment is to elicit a range of possible values for key factors from genebank managers. For instances, a statistical distribution of annual costs per accession, or in-perpetuity cost of conserving all accessions, could be generated based on elicited values.

The @Risk™ software can be used to define or adjust distributions to available data and to perform the sensitivity analysis. The software allows for the substitution of single point values with a probability distribution. A triangular distribution is the simplest distribution to elicit that approximates a normal distribution. This distribution is widely used in decision theory, especially when no sample data are available (Hardaker et al. 1997). The parameters defining the distribution are lowest, highest and most common value. Means, variances and coefficients of variation are easily tabulated from these three values, and repeated sampling from the distributions can be used to generate overall distributions.

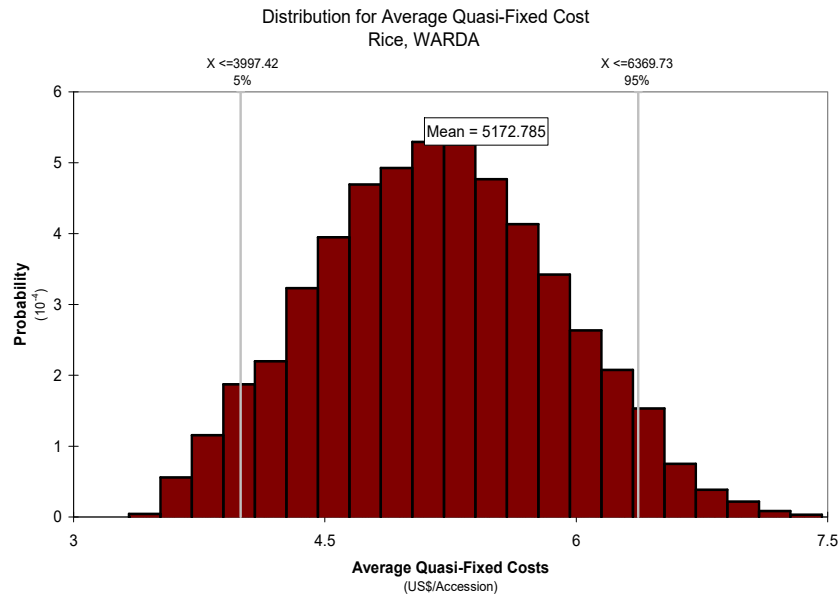
For instance, let us take the number-of-accessions-regenerated-per-year (NREG) as an example of a factor affecting costs in the WARDA genebank. We can ask the genebank manager for his or her “best guess” of the highest, lowest and most common values for NREG conditional on a reference period and technology. Using these three parameters, the software then generates a distribution of values for NREG. We could also generate unconditional distributions across technologies. Instead of a single value for total costs of maintaining a rice accession in the WARDA genebank, we would then have a distribution of values. The software can evaluate the simultaneous effect of more than one factor (input variable) on one or more than one cost variable (output variable). Simulations results using WARDA genebank data are shown in Figures 4a and 4b.

In the decision tool, factors affecting genebank costs are currently included in the “general information” sheet. Although preliminary simulations can be run based on this information, the aim of this task is to evaluate the relationship between performance and

costs. As explained above, costs information is entered by activity. Certain activities must be performed in order to accomplish an operation. It is possible to find one or more performance indicators that are linked to an operation or an activity. The difficult in this case is that some inputs are used and activities performed for more than one operation. For example, regeneration is needed for conservation but also to ensure that genetic materials will be available for users (see Table 4).

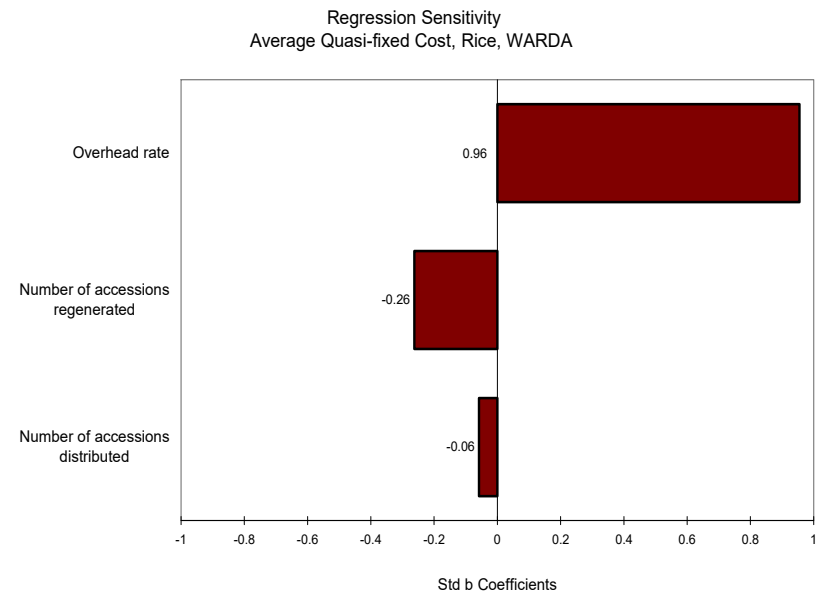
Indices of genetic erosion might be used as performance indicators for the regeneration activity. Genebanks make use of variable, fixed and quasi-fixed inputs to regenerate the material and most importantly in order to maintain a low index of genetic erosion. If the genebank is not performing well and genetic erosion is high (or higher than the standard level/best practice recommendation), how should the manager allocate inputs in order to reduce the index of genetic erosion? Increasing a technician's time in order to regenerate wild rice will most likely reduce the index of genetic erosion, but by how much? Thus, both the effect of input use on performance and the effect of performance on costs are difficult to grasp.

Figure 4a



Instead of reporting just one value this report shows a distribution of values for average quasi-fixed cost of keeping one accession of rice in the WARDA genebank. This distribution of values respond to changes in: 1) number of accessions regenerated, 2) number of accessions distributed, 3) overhead rates, and 4) interval period for regeneration. Each of these variables has been replaced by a distribution of values. With this tool we can measure the effect of one or of multiple variables at the same time.

Figure 4b



This figure depicts the effect of individual distributions over the distribution of average quasi-fixed costs. From the factors evaluated (number of accessions regenerated, number of accessions distributed, overhead rate, and interval period to regeneration) overhead rate has the strongest effect on the increase of average quasi-fixed cost of rice accessions in the WARDA genebank. Average quasi-fixed costs decrease with an increase of number of accessions regenerated or distributed.

Table 4. Relating genebank performance to costs

Goal	Specific Objective ⁸	Activity ⁹	Performance Indicators ¹⁰	Cost Components
Conservation	Coverage of Genepool	- Acquisition	- Average acquisition costs - Number of collected accession per acquisition	- CC: seed health testing facility, greenhouse, lab/office equipment - QFC: scientific and technical staff - VC: labor for seed health testing, introductory planting, seed handling
		- Security duplication /	- Number of accessions safely duplicated	- CC: - QFC: - VC: labor for packing and shipping
	Ensuring security	- Long term storage (collection)	- Seed longevity - Number of accession stored over time	- CC: storage facility, storage equipment, back up power system, storage container - QFC: scientific and technical staff - VC: labor for storage
		- Germination testing and seed health testing	- Germination levels and deviation - Seed quality	- CC: germination testing facility, germination chamber, etc - QFC: scientific and technical staff - VC: labor for germination testing
	Maintaining genetic integrity	- Regeneration	- Number of accessions lost - Genetic erosion	- CC: screenhouse, storage, vernalizer, seed-cleaning equipment, seed-drying equipment, seed-processing facility, seed processing equipment, etc - QFC: scientific and technical staff - VC: labor for regeneration
Use	Ensuring availability	- Medium term storage (active collection)	- Seed longevity - Seed quality	- CC: storage facility, storage equipment, backup power system, seed container - QFC: scientific and technical staff - VC: labor for medium term storage
		- Germination testing and seed health testing	- Germination levels and deviation - Seed quality	- CC: germination testing facility, germination chamber, etc - QFC: scientific and technical staff - VC: labor for germination testing
		- Regeneration	- Number of accessions lost (depending on the crop and conservation method) - Genetic drift indicator	- CC: screenhouse, storage, vernalizer, seed-cleaning equipment, seed-drying equipment, seed-processing facility, seed processing equipment, etc - QFC: scientific and technical staff - VC: labor for regeneration
	Distribution of material	- Distribution (Shipment)	- Number of accessions distributed per year	- CC: seed health testing facility, greenhouse, lab/office equipment, Jacuzzi equipment, etc - QFC: scientific and technical staff - VC: labor for seed health testing duplication, packing and shipping
	Providing information	- Data / Information management	- User appraisal - Number of visitors to website	- CC: general facilities - QFC: scientific and technical staff - VC: labor for database management, catalogue management, etc.

⁸ According to Common Analysis framework of the GPG2 activities

⁹ Following Koo et al. (2004) classification/grouping of activities

¹⁰ Examples from the performance indicators of the CGN in Netherlands

5. Costs Function

What would be the costs of reaching a “best practice” in the CG genebank system? How do location-specific variables affect genebanks costs? Genebank management decisions and their costs implications need to be evaluated to improve the performance of individual genebanks but also of the system as a whole. The second, longer-term objective the decision support tool is to provide answers to questions about the global genebank system. When enough observations have been assembled through application of the decision support tool, a genebank cost function can be estimated and specific hypotheses tested.

Genebanks costs depend on several factors: biological characteristics of the crop conserved, conservation methodology used (in vitro, field germplasm banks), institutional differences (wage structure, cost-sharing opportunities), local climate (for regeneration for instance, general state of infrastructure). The use of econometric methods will permit SGRP to evaluate the system as a whole by disaggregating the effect of the different factors and performance grade on costs. Once these effects have been taken into account in a multivariate regression, it will be feasible to draw conclusions across centers and genetic materials.

Cost function approaches have been used to model other public goods like hospitals and libraries, and this literature can provide insights into how we might specify genebanks costs. Finch and Christianson (1981) modeled the costs function of rural hospitals in US. The purpose of their study was to supply information about hospital costs that be used in making decisions regarding how the provision of health care to rural populations. The authors used quadratic and logistic specifications. The main advantage of the quadratic U-shape function is that a cost minimum can be determined given a fixed level of output. The logistic L-shape cost function implies that costs are decreasing but not in a constant rate to output, similar to what we assumed for genebanks. An additional contribution of this study is the use of output indicators to account for short run and long run costs. Conservation and distribution of genetic materials and information fit this characterization well.

Liu (2002) modeled a cost function for academic research libraries, taking into account the multi-product and multi-service nature of information provision by such organizations, which is comparable to genebanks. The author used a log-linear function and considered that some economies of scale exist in the operation of research libraries. Translog cost models have been also used for evaluating the costs function of research libraries. The use of a translog cost function is convenient when the goal is to determine elasticities of substitution among different inputs. De Boer (1992) used a translog cost function to examine economies of scale and input substitution elasticities of 194 Indiana public libraries.

If the objective of modeling genebank costs is to evaluate the relationship between cost and outputs with current technology and practices, logistic or quadratic specifications suit the purpose. If the objective is however to determine a technical relationship among inputs and outputs, a translog model would do better. It is always possible to test different specifications and evaluate which one adjusts better to our needs.

In specifying the model it is also important to define which cost to model: total costs or average (variable) cost. Modeling total costs would provide some additional information on capital and quasi-fixed costs. Since it is expected that changes in technology (mainly capital inputs) would improve performance of specific activities and operations, it may be better to model total costs than average variable costs. It is also possible to model total cost per operation.

Exogenous variables will include performance indicators for each output, a vector of crop characteristics, a vector of genebank characteristics, and a vector of staff characteristics. Genebank objectives can be classified according to either short or long term goals. This classification would help to determine short- and long-term minimum costs. Outputs related to conservation (covering genepool, maintaining genetic integrity and ensuring security) can be considered of a long-term nature while outputs of germplasm use (ensuring availability, providing information and germplasm distribution) are considered to be short-term outputs.

The general function could then be specified as:

$$TC = f(PI, Cr, Gb, St, L)$$

Where

PI = performance indicators (for either short or long term goals),
Cr = crop characteristics (multiplication strategy, fertilization, etc),
Gb = genebank characteristics (facilities, equipment),
St = staff characteristics (number, skills); and
L = location (dummies).

The selection of appropriate performance indicators is crucial. Conservation indicators might include the number of accessions stored, or indicators of diversity represented by the accessions. Indicators for use of germplasm might be the number of accessions added per year, the number of accession distributed, or the number of users of the genebank. Notice that the number of accessions added reflects the performance of genebank (technical operations) directly, while the other two indicators are related more directly to users and might be more appropriate if the goal is to measure the impact of a genebank.

The vector of crop characteristics includes the type of fertilization (open pollination, cross pollination) and type of seed reproduction system (sexual, clonal, tissue culture, etc). Differences in crops and reproduction system have a definitive effect on the costs. Genebank characteristics that can influence the cost function are related to the type of equipment and facilities. This information is also valuable to determine if the genebank is operating under excess capacity or not, and thus where economies of scale might be achieved. Staff characteristics variables can help to explain the effect on costs of staff qualification and the number of staff working in the genebank. The use of dummy variables is recommended to factor out location specific effects.

6. Considerations and next steps

- The decision support tool is in an early stage of development. The framework used to develop the tool (activities, operations, objectives, etc) is based on research produced during the first phase of the GPG project. The goal of this work is to produce a flexible tool that can be adjusted to represent a range of operations and genetic materials.

- A consensus regarding a set of simple, quantifiable performance indicators is central to further progress in developing the tool. To be useful, the tool must be modified in order to incorporate data that measure performance.
- The tool can be used to produce annual costs reports and a sensitivity analysis based on simulations. These evaluations can be accomplished per genetic material in a specific genebank. The longer term goal, however, is to evaluate genebank performance for the global genebank system. The tool can be used to assemble relevant data, and based on a review of cost studies of provision of public goods, we recommend econometric analysis as a means of evaluating the system.
- The next steps involved in implementing the decision support tool are: a) assemble input use, costs information and feedback from genebank managers, b) add an input sheet for entering performance indicators, c) examine, with genebank managers, how costs can be structurally linked to performance indicators.

References

- Altoveros, N. C., and R. V. Rao. 1998. Analysis of information on seed germplasm regeneration practices. In *Regeneration of Seed Crops and their Wild Relatives: Proceedings of a Consultation Meeting, 4-7 December 1995, ICRISAT, Hyderabad, India*, ed. J. M. M. Engels, and R. R. Rao. Rome: IPGRI.
- Burstin, J., M. Lefort, M. Mitteau, A. Sontot, and J. Guiard. 1997. Towards the assessment of the costs of genebank management: conservation, regeneration, and characterization. *Plant Varieties and Seeds* 10: 163-172.
- Calles, T., M. E. Dulloo, J. M. M. Engels, and I. van den Houwe. 2007. *Best practices for germplasm management: A new approach for achieving genebank standards*. Bioversity International.
- Day-Rubenstein, K., P. Heisey, R. Shoemaker, J. Sullivan, and G. Frisvold. 2005. *Crop genetic resources: An economic appraisal. A report from the economic research service*. Economic Information Bulletin Number 2. United States Department of Agriculture.
- De Boer, L. 1992. Economies of scale and input substitution in public libraries. *Journal of Urban Economics* 32 (2): 257-268.
- Engels, J. M. M., and R. R. Rao. 1998. *Regeneration of Seed Crops and their Wild Relatives*. Rome: IPGRI.
- Finch, L., and J. B. Christianson. 1981. Rural hospitals costs: an analysis with policy implications. *Public Health Reports* 96: 423-433.
- Fowler, C., and T. Hodgkin. 2004. *Plant genetic resources for food and agriculture: Assessing global availability*. Vol.2004. 29 ed.
- Hardaker, J. B., R. B. M. Huirne, and J. R. Anderson. 1997. *Coping with risk in agriculture*. Wallingford: CAB International.
- Koo, B., P. G. Pardey, and B. D. Wright. 2003. The economic costs of conserving genetic resources at the CGIAR centres. *Agricultural Economics* 29: 287-297.
- . 2004. *Saving seeds: The economics of conserving crop genetic resources ex-situ in the future harvest centres of the CGIAR*. Oxfordshire: CABI Publishing.
- Koo, B., and B. D. Wright. 2008. The optimal timing of evaluation genebank accessions and the effects of biotechnology. *American Journal of Agricultural Economics* 82 (4): 797-811.
- Liu, L. G. 2002. The cost function and scale economies in academic research libraries. *Library Trends* 2002: 406-420.

- Pardey, P. G., B. Koo, M. E. Van Dusen, B. Skovmand, and S. Taba. 2001. Costing the Conservation of Genetic Resources: CIMMYT's Ex Situ Maize and Wheat Collection. *Crop Science* 41: 1286-1299.
- Rao, N. K., J. Hanson, M. E. Dooloo, K. Ghosh, D. Nowell, and M. Larinde. 2006. *Manual of seed handling in genebanks*. Handbook for Genebanks No. 8, Maccarese, Italy: Bioversity International.
- Smale, M., and A. G. Drucker. 2007. Agricultural Development and the Diversity of Crop and Livestock Genetic Resources: A Review of the Economics Literature. In *Biodiversity Economics*, ed. A. Kontoleon, U. Pascual, and T. Swanson. Cambridge: Cambridge University Press.
- Smale, M., and B. Koo. 2003. *Genetic Resources Policies: What is a Gene Bank Worth?* IFPRI: Washington, D.C.: IFPRI, IPGRI, and the Systemwide Genetic Resources Program.
- Taba, S., D. van Ginkel, D. Hoisington, and D. Poland. 2004. *Wellhausen-Anderson Plant Genetic Resources Center: Operations Manual*. El Batan, Mexico: CIMMYT.
- Virchow, D. 2003. Current expenditures on crop genetic resources activities. In *Efficient conservation of crop genetic diversity*, ed. D. Virchow. Berlin: Springer-Verlag.
- Virchow, D. 1999. *Spending on conservation of plant genetic resources for food and agriculture: How much and How efficient?* ZEF Discussion Papers on Development Policy. Bonn: Center for Development Research.

Annex 1: Preliminary Performance Indicators for the GPG Project
(after the meeting in October 2007 in Lunteren)

A) Technical Group

Coverage of Diversity (Gap)

- Geographic and/or taxonomic coverage of collection assessed? Y/N

Risk Management

- Do you have an HR plan? Y/N
- Risk management plan in place? Y/N
- % safety duplication?

Quality

- 1) Quality management in place? Y/N
- 2) To what extent (%) are you up-to-date with the monitoring of seed viability?
- 3) To what extent (%) are you up-to-date with the regeneration?
- 4) Number of accessions lost since last year?
- 5) Do you have a strategy to maintain the genetic integrity? – to be worked on

Distribution

- 1) % of material free of targeted pathogen?
- 2) % Customer satisfaction responses
- 3) Cost recovery plan Y/N? – to be worked on
- 4) Core collection characterized Y/N? – to be worked on

Regeneration

- % of accessions with passport data? – to be worked on
- that have essential passport and identity information available on the web (for baseline, comparative purposes, morphology, herbarium specimens, molecular characterization, photos, seed files, etc.) – to be worked on
- % of deviants in multiplied materials - out

B) Impact and relevance

Green = OK

Blue = to be worked on

Red = out

1. Number & diversity of users

- Number of users

- Diversity of users (universities, NARS, private, CSO, CGIAR, other genebanks)
- Number of countries distributed to
- Number of 'new' countries compared to past 4 years

- Number of 'new' users compared to past 4 years –to be worked on
- reference number cited in scientific publications –to be worked on
- reasons for requesting materials - out

2. Number of samples

- Number of samples distributed
- Number of distinct accessions

- Number of accessions, distinct from past 4 years –to be worked on
- Number of focused requests (how to define 'focused'?) - out

Other indicators

- Average number of data, available on the web: passport, evaluation, characterization data
 - (future: add a measure for user-friendliness of the website)
- # of technical/policy research papers: scientific, popular, . . . (together – without impact factors)
- Is there a marketing & awareness strategy (indicators for relevance to be developed)
- Strategic plan (indicators for relevance to be developed)

- # of people trained –to be worked on
- # of contact moments with policy relevant people (define contact moment) –to be worked on
- Hits on website/database X time –to be worked on

C) International collaboration

How much do we collaborate? Internally (between Centers) among conservers and users

- # joint projects
- # joint outputs (indicator of quality)
- Collective actions
- Participation in annual SGRP meetings
- Best practices
- Germplasm exchange
- # of FTE partners (size & resources committed)

- # successful collective proposals - out

How much do we collaborate with external conservers and users?

- # joint projects
- # joint outputs (indicator of quality)
- # of FTE partners (size & resources committed)
- # joint/participatory planning events

- # complementary/supportive activities –to be worked on
- # *successful collective proposals - out*

Behaviour change

- Ratio of project coordinatorships/total projects
 - participation without control
 - give others ownership
- # of publications with partners (ratio/total) –to be worked on

International Treaty

- # of international exchanges of germplasm –to be worked on
- # SMTAs for non CG collections –to be worked on

Distribution is proactive in demonstrating benefit of exchange

Categories of partners

- # of links (active contacts) to on-farm/community-based partnerships? –to be worked on

Partner/customer satisfaction

- Yearly survey – rating satisfaction 1-5 –to be worked on
- Joint project partners –to be worked on
- Key partner(s) surveyed – i.e. National programs working on same crop – common objective – to be worked on
- How well was service performed? –to be worked on

Note: active partnerships/links to cover gaps in the genepool.

Annex 2: Decision Support Tool

Microsoft Excel - Cost-structure-WARDA GRU-DH.xls

File Edit View Insert Format Tools Data Window @RISK Help

Type a question for help

100% Arial

COST STRUCTURE OF GENE BANK MANAGEMENT

Introduction

This spreadsheet has been designed to calculate the current costs of operating your genebank. If you enter the relevant cost information in the following four sheets, several cost reports regarding your genebank will be automatically calculated in the ensuing sheets. The excel file is composed of 9 spreadsheets. In addition to this "introduction" spreadsheet (blue tag), there are 4 spreadsheets for data entry (white tag), and 4 spreadsheets for pre-defined reports (red tag). This tool can help you understand the cost structure of operating your genebank in the relevant year, and can be used in making strategic decision for the efficient operation of your genebank in the long run.

Cost classification

The costs considered in this spreadsheet can be classified as:

- Capital cost (or fixed cost): building, equipment, furniture, etc.
- Operating cost (or variable cost)
 - Labor cost: genebank manager, scientists, technicians, workers (temporary or permanent)
 - Non-labor costs: electricity, fuel, water, office expenses, etc.

Note: Some of the labor costs (such as genebank manager and scientists) can be considered fixed because it does not directly depend on the variation of operation in the short run. We separately categorized this as quasi-fixed cost in the data entry sheets. If the proportion of the fixed cost is large in the genebank operation, there may be room for consolidating several genebanks in a single facility.

Classification of genebank activities

Each genebank is operated differently depending on the specific environments and it is difficult to incorporate all the differences in a single framework. However, this tool categorizes various genebank operations in the following classification based on Koo et al. (2004). You can adjust this framework by adding or deleting activities.

Activity	Explanation	Code
Acquisition	This may involve the collection activities in the fields or the activities related to receiving and processing newly introduced accessions.	ACQ
Characterization	This is the activity of recording the characteristics of each accession, often conducted during the regeneration process.	CHA
Safety duplication (or security duplication)	This is the activity of sending sample accessions to different location for safety reason (i.e., backup collection).	SDUP
Long term storage	This activity is for the conservation of accessions in the long term storage facility.	LTS

Ready NUM

Annex 3: Decision Support Tool – Summary report

1) Capital Cost

Activities	Crops				
	Rice	Wild rice	0	0	0
Acquisition	270.50	2.82	0.00	0.00	0.00
Characterization	8,672.09	90.54	0.00	0.00	0.00
Safety duplication (or security duplication)	112.37	1.17	0.00	0.00	0.00
Long term storage	112.37	1.17	0.00	0.00	0.00
Medium term storage	385.78	4.03	0.00	0.00	0.00
Germination testing (or viability testing)	4,216.79	44.03	0.00	0.00	0.00
Regeneration (or multiplication)	4,353.57	45.45	0.00	0.00	0.00
Seed health testing	333.06	3.48	0.00	0.00	0.00
Dissemination (or distribution)	342.37	3.57	0.00	0.00	0.00
Information and data management	1,516.51	15.83	0.00	0.00	0.00
General management	0.00	0.00	0.00	0.00	0.00
Other 1 Temporary storage	0.00	0.00	0.00	0.00	0.00
Other 2	0.00	0.00	0.00	0.00	0.00
TOTAL	20,315.42	212.10	0.00	0.00	0.00

2) Quasi Fixed Costs (skilled labor)

Activities	Crops				
	Rice	Wild rice	0	0	0
Acquisition	14,947.85	156.06	0.00	0.00	0.00
Characterization	16,166.65	168.79	0.00	0.00	0.00
Safety duplication (or security duplication)	8,044.49	83.99	0.00	0.00	0.00
Long term storage	6,135.94	64.06	0.00	0.00	0.00
Medium term storage	16,364.58	170.85	0.00	0.00	0.00
Germination testing (or viability testing)	12,773.60	133.36	0.00	0.00	0.00
Regeneration (or multiplication)	28,908.35	7,227.09	0.00	0.00	0.00
Seed health testing	2,919.52	30.48	0.00	0.00	0.00
Dissemination (or distribution)	23,589.15	246.28	0.00	0.00	0.00
Information and data management	33,846.63	353.37	0.00	0.00	0.00
General management	2,137.68	22.32	0.00	0.00	0.00
Other 1 Temporary storage	21,306.03	222.44	0.00	0.00	0.00
Other 2	0.00	0.00	0.00	0.00	0.00
TOTAL	187,140.47	8,879.09	0.00	0.00	0.00

3) Variable Cost - Labor

Activities	Crops				
	Rice	Wild rice	0	0	0
Acquisition	5,621.45	4.07	0.00	0.00	0.00
Characterization	15,702.66	5.58	0.00	0.00	0.00
Safety duplication (or security duplication)	893.41	9.33	0.00	0.00	0.00
Long term storage	254.43	2.66	0.00	0.00	0.00
Medium term storage	330.72	3.45	0.00	0.00	0.00
Germination testing (or viability testing)	8,180.61	36.43	0.00	0.00	0.00
Regeneration (or multiplication)	21,326.17	39.46	0.00	0.00	0.00
Seed health testing	4,691.41	0.00	0.00	0.00	0.00
Dissemination (or distribution)	11,586.19	23.00	0.00	0.00	0.00
Information and data management	34,877.99	315.16	0.00	0.00	0.00
General management	0.00	0.00	0.00	0.00	0.00
Other 1 Temporary storage	203.57	2.13	0.00	0.00	0.00
Other 2	0.00	0.00	0.00	0.00	0.00
TOTAL	103,668.62	441.27	0.00	0.00	0.00

4) Variable Costs - Non Labor

Activities	Crops				
	Rice	Wild rice	0	0	0
Acquisition	0.00	0.00	0.00	0.00	0.00
Characterization	7.44	0.08	0.00	0.00	0.00
Safety duplication (or security duplication)	7.44	0.08	0.00	0.00	0.00
Long term storage	0.00	0.00	0.00	0.00	0.00
Medium term storage	3,195.37	33.36	0.00	0.00	0.00
Germination testing (or viability testing)	74.39	0.78	0.00	0.00	0.00
Regeneration (or multiplication)	13.53	1.50	0.00	0.00	0.00
Seed health testing	7.44	0.08	0.00	0.00	0.00
Dissemination (or distribution)	14.88	0.16	0.00	0.00	0.00
Information and data management	14.88	0.16	0.00	0.00	0.00
General management	0.00	0.00	0.00	0.00	0.00
Other 1 Temporary storage	637.59	6.66	0.00	0.00	0.00
Other 2	0.00	0.00	0.00	0.00	0.00
TOTAL	3,972.94	42.84	0.00	0.00	0.00