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

Daniela Horna

The European Cooperative Program for Plant Genetic Resources (ECPGR) is working on the implementation of a European Genebank Integrated System (AEGIS). The goal of AEGIS is to create an integrated genebank system for conserving the genetically unique and important accessions of Europe and making them available for breeding and research (ECPGR 2008). In addition to other clear advantages of AEGIS, this system is expected to increase, in the long term, the costs effectiveness of the collections management. The framework presented here discusses elements needed for evaluating and furthermore for monitoring this costs effectiveness.

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, and with what mix of inputs. 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. Most of the benefits of genebank collections are public goods whose values are both expensive to estimate and likely to be unreliably 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 estimate of the corresponding benefits, for many decisions, it may not be necessary to estimate benefits.

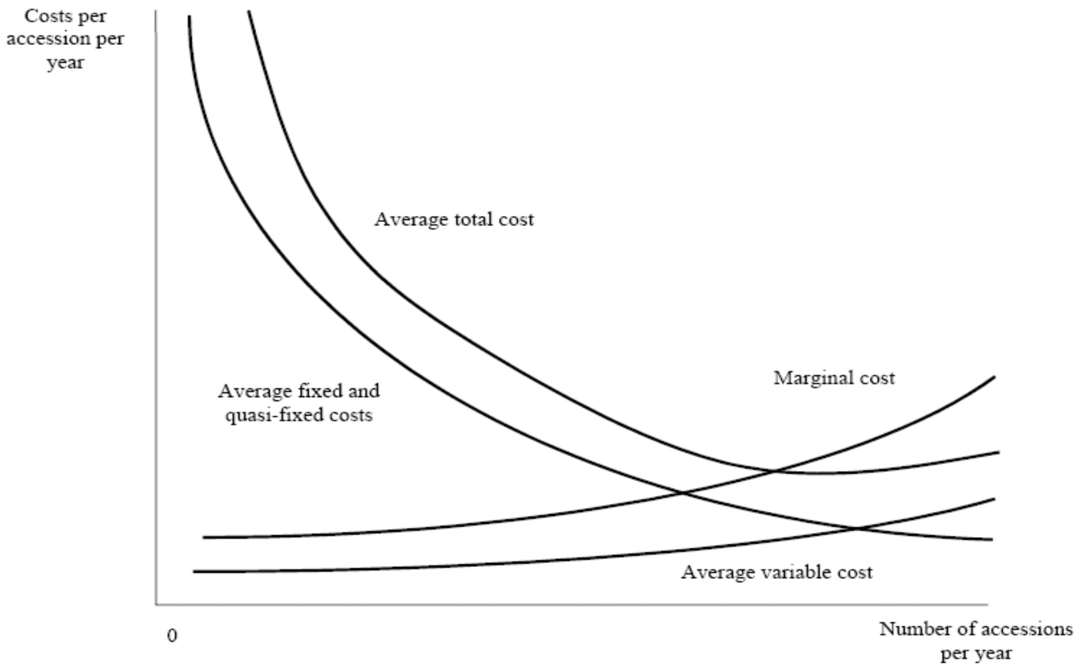
A few studies have addressed the issue of cost implication of germplasm conservation, but without an explicit micro-economic framework. Virchow (2003; 1999) used surveys to collect national conservation expenditure for 39 countries and estimated per-accession cost of annual conservation for each country. 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.

Analysis of Genebank Costs: Costs Minimization Framework

Before introducing the framework a few economic concepts need to be described in the context of a genebank. For a genebank, average costs are the costs of managing one accession. Marginal costs, on the other hand, 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 bound estimates of the marginal costs. Figure 1 illustrates how average and marginal costs are thought to change with amounts produced (for example, the number of seeds stored, regenerated, disseminated, etc).

Figure 1: Genebank average and marginal cost



Source: Pardey et al. 2001.

Data Collection: Decision Support Tool

To evaluate the costs effectiveness of a genebank or system of genebanks it is ideal to count with time series data. This information must be gathered in a uniform way in order to facilitate comparison across genetic materials and across centers. Unfortunately, genebanks often do not keep systematic records of their operational costs. Data collection then requires a tool that can 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.

In order to collect the cost information for genebanks, 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, 4 input sheets (non-labor variable inputs, variable labor, quasi-fixed labor and capital inputs) and 4 output sheets or reports. 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, non-labor variable, labor variable, and quasi-fixed costs. In addition to a summary overview by crop and input costs, three kinds of output reports can be generated according to genebank managers needs. The most basic report can show costs per input category, genetic material, and operation, as shown in Table 1. This kind of report provides information about both total costs and average costs per accession. A graphic representation of the distribution of total costs can be easily developed (Figure 2). Genebank managers are interested in knowing not only how much it costs to maintain the current level of operations but also how much it would costs to keep all these accessions in perpetuity. This information can be easily generated once the basic cost information is gathered. Such information is useful when justifying genebank funding or investment in ex-situ conservation.

Sensitivity Analysis with Simulations

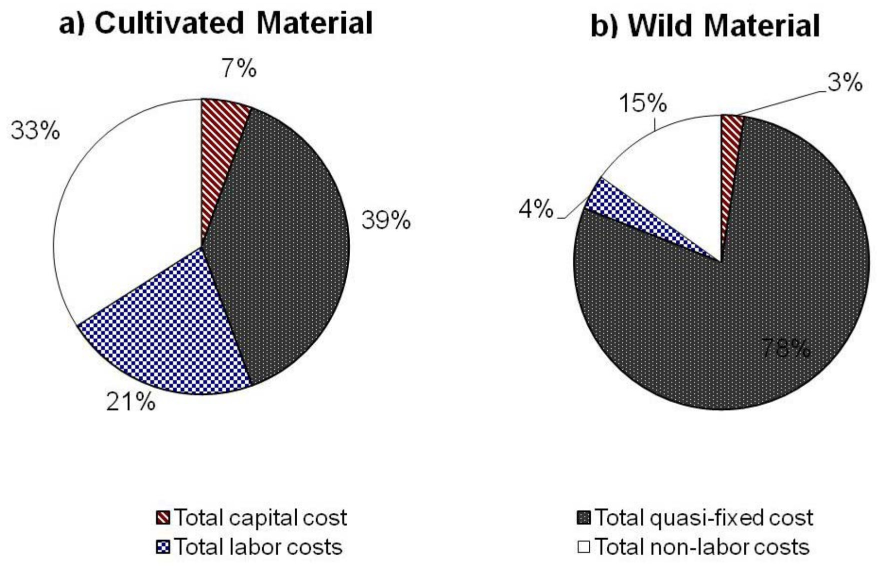
While the reports help to understand the structure of genebank costs and their distribution across operations, objectives and over time, nothing can be inferred about the factors that affect these costs. For this reason, although it is possible to compare reports across genebanks, the reports are not a picture that enables tackling strategic decisions. A feasible way to extend the use of the tool is using simulations.

Table 1. Genebank Report, Total and Average Costs

Genetic Material 1

Activities	Number of accessions	Total capital cost ($/year)	Total quasi-fixed cost ($/year)	Total labor costs ($/year)	Total non-labor costs ($/year)	Average capital cost ($ /acc)	Average quasi-fixed cost ($ /acc)	Average variable cost ($ /acc)	Total AC ($ /acc)
Acquisition	4,950	3,408.51	17,235.23	5,167.51	3,524.80	0.69	3.48	1.76	5.93
Characterization	2,000	1,704.26	10,359.51	3,012.34	1,671.48	0.85	5.18	2.34	8.37
Safety duplication	9,450	1,704.26	5,602.95	1,474.33	4,247.28	0.18	0.59	0.61	1.38
Long term storage	83,930	26,048.57	6,074.64	3,012.34	9,893.62	0.31	0.07	0.15	0.54
Medium term storage	86,080	36,016.41	15,393.79	5,363.43	15,977.30	0.42	0.18	0.25	0.85
Germination testing	29,250	22,724.62	10,734.21	7,249.21	3,783.79	0.78	0.37	0.38	1.52
Regeneration	7,400	87,899.06	33,064.54	125,118.99	8,213.50	11.88	4.47	18.02	34.36
Seed health testing	7,300	1,449.36	3,100.81	300.01	0.00	0.20	0.42	0.04	0.66
Distribution	6,200	7,804.44	21,624.13	4,889.11	9,151.05	1.26	3.49	2.26	7.01
Data management	86,800	1,704.26	31,617.32	6,858.03	6,773.35	0.02	0.36	0.16	0.54
Total	**N.A.**	**190,463.75**	**154,807.12**	**162,445.30**	**63,236.18**	**16.58**	**18.62**	**25.96**	**61.16**

Figure 2. Graphic representation of the distribution of genebank costs



When it is augmented by sensitivity analysis and simulations, the information collected with 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 instance, 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 a 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 an accession in the 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).

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.

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 a wild material 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.

Next steps

The decision support tool is in an early stage of development. The goal of this work is to produce a flexible tool that can be adjusted to represent a range of operations and genetic materials. 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 the performance of a 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, analyze the system.

The tool has been developed to be implemented for individual genebanks. Comparisons across genebanks are limited by the different conditions under which the genebank operate (location, technical skill of the labor, budget, etc.). AEGIS is a complex system of genebanks working with a large number of genetic materials. Evaluating the costs effectiveness of an individual genebank is a relatively straightforward exercise. In this case, it is common to assume that genebank managers are optimally allocating resources and at the same time minimizing costs. Evaluating the effectiveness of a system of genebanks is a much more challenging exercise given that there are other considerations that are as equally important as the costs minimization goal. These considerations are often more difficult to value. For example, a particular genebank may be having very high conservation costs, but it is more important for the system to count with the participation of this genebank. What is then the value that we have to assign to "participation in AEGIS." What other considerations are there to pay attention to? How can we incorporate these other considerations to our current framework?

In order to answer the questions above an option would be to collect costs information for a system of genebanks currently functional like EURALLIVEG. This exercise can give us a better idea of how AEGIS could work and at the same time can help us to develop additional components in the tool that can address the system needs. One constraint with this approach is that the genebanks members of EURALLIVEG can be self selected and could not be a representative sample of AEGIS.

An alternative would be to collect costs information from a random selection of AEGIS genebanks that work with the model crops (avena, prunus, brassica and allium). The genebanks could be previously classified according to predetermined criteria, size of operation, location, or level of specialization. The random selection would allow arriving to recommendations and answers to management questions for the individual genebanks and for the system. This is however a more resource intensive approach.

Final Remark

In the transformation process from a genebank working individually to a more integrated system of genebanks, it is essential to have an idea of the costs involved in genebank operations. It is even more valuable however to have periodic information that can be analyzed overtime. In the same way that germplasm information is relevant to decide what to conserve and how to do it, the monitoring of costs information is crucial to address these key management questions.

References

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.

ECPGR. 2008. A strategic framework for the implementation of a European genebank integrated system (AEGIS). *Discussion Paper.European Cooperative Programme for Plant Genetic Resources (ECPGR)* .

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. 2004. *Saving seeds: The economics of conserving crop genetic resources ex-situ in the future harvest centres of the CGIAR.* Oxfordshire: CABI Publishing.

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.