Weighting of Agricultural Research Results:
Strength and Limitations of the Analytic Hierarchy Process (AHP)

Frank Hartwich

Universität Hohenheim, Institut für Agarökonomie und Soziologie in den Tropen und Subtropen, Lehrstuhl für Marktlehre in den Tropen und Subtropen, e-mail: hartwich@uni-hohenheim.de

Introduction

The inherent complexity and uncertainty surrounding many decisions in management and policy have led to the development of a large range of tools to support decisions. Some of those tools involve in attaching weights and deriving priorities to distinct options and choices. The Analytic Hierarchy Process (AHP) is a recent addition to the set of methodologies for assessing and allocating weights and priorities. AHP is a decision analysis technique that uses judgements form a group of relevant decision makers along with hierarchical decomposition of a problem to derive a set of ratio-scaled measures for decision alternatives. With the AHP the analyst structures a problem hierarchically and then, through an associated measurement-and-decomposition process, determines the relative priorities consistent with overall objectives. AHP has been developed initially by Thomas L. Saaty at the University of Pittsburgh, Pennsylvania and found its way into management and operations research, marketing studies and various decision support systems in the private sector and in government policy. The recent development of the user friendly AHP decision support software, Expert Choice ©, gave rise to its popularity.

Management of agricultural research and development (research) constitutes a field where crude difficulties exist in attaching prices, weights or priorities to respective research results. Research results may be very different in their character and relevance. Also, the research process by itself is deriving unforeseeable results. Different approaches have been applied in the evaluation of agricultural research in order to derive priorities or weights to research results. They range from very applied approaches as for example scoring methods to methodological complex approaches involving economic surplus measurement. At ISNAR Braunschweig (1998) recently developed an ex-ante evaluation method for priority setting in biotechnology research which is based on the judgements of peers applying the Analytic Hierarchy Process (AHP). This method was particularly successful in accounting for impact with the use of subjective peer knowledge. Also at ISNAR this method was applied to ex-post evaluation of agricultural research. First steps involved the development of a cumulative performance indicator (CPI) to describe the performance of research institutions (Hartwich, 1998). This approach combined quantitative data on research outputs with weighted scores based on qualitative assessments of research using AHP. Analysis of an institution's research performance yield a weighted outputs indicator, which can be divided by the institution's budget or other input measure to be transformed into a performance indicator. In the further development of the methodology AHP was combined with the method of Data Envelopment Analysis (DEA). Within categories the AHP was used for weighting different outputs. Then DEA was applied to avoid further agglomeration of categories calculating Pareto efficiencies. This led to an overall weighted technical efficiency indicator with which the research units under investigation could be compared (Hartwich, forthcoming).

In this paper the usefulness of AHP in weighting agricultural research results in ex-post research evaluation is discussed. As a basis for a theoretical discussion both a review of existing methods which are potentially able to derive weights or priorities and an in-depth description of the method giving a practical example are presented. In the first section the principal existing methods for deriving weights and priorities in agricultural research are reviewed. In section two, the background of AHP is briefly described. In section three the principle steps and algorithms of the Analytic Hierarchy Process methodology are presented. Respective procedures involved are illustrated using an example of a decision problem drawn from the field of agricultural research evaluation. In section four strength and limitations of the AHP are discussed and its particularly usefulness in ex-post agricultural research evaluation is investigated before concluding.

1. Research Evaluation and Weighting of Multiple Research Outputs

Research evaluation is judging, appraising, or determining the worth, value , or quality of research, whether it is proposed, on-going, or completed. This is done in terms of its effectiveness, efficiency, impact, and relevance (Horton et al, 1993). Effectiveness refers to the degree to which goals have been achieved. Efficiency refers to the cost-effectiveness of activities. Impact refers to the broad, long-term effects of research. And relevance refers to the appropriateness and importance of goals and objectives in relation to assessed needs.

Assuming that a research operation to be evaluated results in a set of indicators of very different research outputs which cannot be compared. This is actually true for agricultural research where the results may be related to specific technologies, to general knowledge which can be used by farmers and extensionists, to knowledge to be used by further research and other research organizations, or to knowledge used in education. Different evaluation approaches exist which try to integrate diverse performance indicators in order to improve the analytic capability of the evaluation. When involving in an aggregation of outputs the need arises to weigh or prioritize different research outputs.

1.1 Existing Methods for Identification of Weights and Priorities in Agricultural Research

Different ways of weighting and prioritizing research can be found in the literature. Mills and Kamau (1998) for example distinguish four most common methods of priority setting in agricultural research: Congruency, cost benefit, economic surplus, and scoring. Krawiek (1984) referring to research project priority setting in the engineering sector presents a somewhat different categorization of methods, i.e.: economic, decision theory (which clarifies trade-offs and risks of a research operation with the use of simulations), constrained optimization (which involves in optimization of a portfolio of projects with the use of linear programming), and scoring. Falconi (1993) reviewing economic techniques distinguishes four methods of research evaluation: economic surplus approach, econometric methods, mathematical programming and multiple criteria models. Drawing from the above categorization which all distinguish evaluation methods by the criteria of disciplines five general classes of methods for identification of weights and priorities in agricultural research can be distinguished (see Table 1).

(1) Econometric Methods

Econometric methods (mostly regression analysis) can be used to estimate agricultural production, supply, cost or profit functions. Econometric methods are used to estimate changes in output resulting from the increased use of conventional inputs (such as fertilizers, pesticides, labor, and machinery) and a range of other factors as for example the extent to which agricultural research is provided. Investment in research, either in human resources or in equipment and activities, can be included as an independent variable in the functions. The estimated coefficient on the research variable can then be used to calculate the marginal product of research and the marginal rate of return to research investments. Those could be used as scores to derive priorities or weights for the technology under investigation.

Econometric methods are more accurate than other methods at assessing the contribution of past research to changes in total output. Provided that reliable historical data are available they are useful in the estimation of spillovers of research benefits among different zones and commodities, interactions between the public and private sectors, and the effects of economic and trade policy (Falconi, 1993). However, the method is not useful when a range of different research operations are compared for the purpose of priority setting or performance evaluation because: (1) Inputs in research have to be cut to those portions, which have been actually used in the respective research operation. Costs of common use of resources, e.g. laboratories, are difficult to split. (2) Only those benefits have to be include which are the result of the research operation. This is only feasible on very aggregated levels, for example the productivity increase in maize resulting from all research on maize. It would be hardly possible to distinguish between results from maize breeding research and research on maize agronomy. (3) The method is of little use when ranking nonprodcution-oriented research areas, like social science and economic research or natural resource management research.

(2) Budgeting Methods

Budgeting methods involve economic concepts such as discounted cash flow, net present value, and interest rates of return to relate the actual costs to the selling value of the research result (e.g. market price for a patent or a technology). The value of each research result to be compared is calculated separately. The thus derived economic measures can then be applied to the weighting or priority identification process. The capital budgeting techniques are well defined and procedures for applying them are easily found in common practice (Krawiec, 1984). However, their practical applicability in the identification of weights or priorities in agricultural research is limited because: (1) market prices for research results from public agricultural research (e.g. patents and technologies) are hardly available, (2) The method does not account for social benefits as market prices of technologies and patents do not reflect social prices and do not include externalities, (3) Inputs in research have to be cut to those portions which have been actually used in the respective research operation. Costs of common use of resources, e.g. laboratories, are difficult to split, and (4) more sophisticated methods are required when multiple funding levels are involved.

(3) Economic Surplus Methods

This method is an extension of budgeting methods. Evaluation of economic surplus use quantitative methods to assess the costs and benefits of agricultural research activities in terms of real costs and benefits to society. The economic-surplus approach estimates returns to investment by estimating, in a first step, the benefits from research in terms of the change in consumer and producer surpluses that result from technological change. Often production, cost or profit functions are estimated for this purpose describing the benefits in production, cost reduction or profit extension of the new technologies developed. In a second step estimated economic surplus is brought together with research costs to estimate an internal rate of return. Economic surplus methods can be used to weight and prioritize research as well as for other purposes of assessing research impact, analyzing change in agricultural productivity, estimate distribution of research benefits among different groups, calculate return on investments of government spending, assess the spillover of research benefits among different technologies, commodities, regions, or countries, or estimate the distribution of research benefits among producers and consumers (Falconi, 1993). Economic surplus is mostly used to assess long-term results at the user end of research and mostly only one research result or technology, e.g. a new seed variety, is analyzed. Results often depend on a wide range of exogenous and environmental factors. The economic surplus approach has been widely used to calculate benefit-cost ratios, internal rates to return, and net present values of benefits from agricultural research. In combination with econometric methods the economic surplus analysis constitutes a sophisticated approach to agricultural research evaluation being able to account for a large set of possible environmental effects. However, the disadvantages in the application to the identification of weights and priorities in research are: (1) Inputs in research have to be cut to those portions which have been actually used in the respective research operation. Costs of common use of resources, e.g. laboratories, are difficult to split. (2) It involves substantial costs for collecting, processing, and interpreting economic and technical data. (3) It is not well suited to ranking non-commodity research operation, such as socio-economic research, natural resource management research, and interdisciplinary and basic research. (4) By itself it cannot account for multiple outputs. This can only be achieved when scoring or programming methods are included.

(4) Programming Methods

Programming methods based on linear and dynamic programming techniques can be applied to maximize the net profits to a research unit (a research program, an institute, an organization) by selecting an optimum portfolio of projects or research operations. The method identifies optimal portfolios based on multiple goals and resource constraints. Like the scoring approach, mathematical programming uses weights to reflect the political and economic importance of a set of research objectives. Mathematical programming methods for establishing weights and priorities in research have hardly been used in practice because of the inherent uncertainty of research (Krawiec, 1984). The limitations of programming methods for research project selection and resource allocation are: (1) They cannot be applied to a situation where the research is more fundamental in nature, projects are less well defined, and the benefits may not be quantifiable (Bell and Read, 1979). (2) They only can be applied to a defined set of common research activities within one administrative unit which has the mandate to allocate the use of resources, e.g. within the program. Comparison among different research units which cannot be put in one programming set (e.g. due to different administrative affiliations) is not possible. (3) They encounter difficulties in the formulation of the constraints matrix.

Somewhat related to the above programming methods is Data Envelopment Analysis (DEA). This method enables frontier estimation with the use of non-parametric programming models leading to a ranking of all units of observation on the basis of technical efficiency scores. The focus is not on the estimation of an average technology production function used by all units but to identify the best practicing units, construct a best practice production frontier and relate all units of analysis to this frontier. DEA uses the principle of Pareto efficiency to determine relative efficiency of different units. The distinct character of DEA is that when pulling different research outputs (results) together no weighting is required. The relation between research results is rather analyzed under Pareto efficiency in a linear programming model allowing that different outputs have different dimensions. Limitations of DEA are: (1) The estimated production frontiers have no statistical properties and tests to be evaluated upon, (2) non-parametric frontiers attribute all deviations from the frontier to inefficiency and do not account for random influences. Hence, they are particularly sensitive to outliers and measurement errors. (3) In cases where a common production technology exists (as it cannot be assumed among agricultural research units) the application of frontier approaches is not useful. In parametric approaches the usefulness of applying a frontier model rather than an average practice model can be tested. Since DEA is a non-parametric technique, statistical hypothesis tests are difficult and are the focus of ongoing research. However, efficiency scores close to one will indicate that it is rather an average production technology which represents the production conditions in the sample of units of analysis. DEA than has little discriminatory power and hardly any conclusions can be drawn because all units produce under the same conditions. (4) Technical efficiency scores derived from frontier estimation vary according to sample sizes. In very small samples, efficiency scores seem to be very high and DEA will thus have little discriminatory power. (5) Ranks from different DEA studies cannot be compared.

(5) Multiple Criteria Methods

Multiple criteria methods (or weighted-criteria or scoring methods) involve the identification and weighting of several criteria to allow research topics to be ranked on the basis of a composite score. Multiple criteria methods use either qualitative or quantitative criteria. If quantitative data is used, e.g. the production value of a set of different crops in a region, weights can be applied in congruence to those quantities (congruency approach). In qualitative multiple criteria methods, peers from science and technology rank the relative importance of a set of performance indicators with regard to their contribution to the overall goal of the research unit. The ranking is translated into weights with which the value of the indicator is multiplied. The method is usually extended to research that is part of one overall goal but which has a set of different objectives. Weights will be assigned to each objective (reflecting their relative importance) and these weights will be used to calculate a single measure of the overall contribution of each indicator to a combined set of objectives.

scoring methods, which are relatively easy to understand and apply, are specifically designed to incorporate non-economic criteria. Multiple criteria methods are particularly useful when dealing with widely different types of objectives. Norton (1993) sees them as appropriate tools in research priority setting when data, time, and analytical capacity is limited. Krawiec (1984) argues, that multiple criteria methods in research evaluation are highly useful to provide decision makers with a target. The limitations of multiple criteria methods are: (1) It is easy to misuse multiple criteria methods when biased subjective views or poorly measured criteria are included. (2) Criteria for the identification of weights or priorities may be overlapping. And (3) The scores derived often fail to discount future benefits and costs. Social benefits and costs are estimated by a set of possibly biased peers (Norton, 1993).

The AHP method as described in this paper can be subsumed under multiple criteria methods. It is a further development of simpler scoring methods as it provides rigid hierarchical structuring of the weighting problem and applies a particular algorithm in the weighting process which makes it less subjective. The Delphi method can be seen as another extension of multiple criteria techniques. The Delphi method is used to derive weights from a group of experts which judge alternatives each with respect to their personal criteria. Often the method is used for prognosis. It can be, however, used to derive weights for future or past research operations. The method applies weighting through peers in successive rounds through questionnaires. After the first round the weights of the experts are aggregated, usually through the computation of the mean. In the next round the participating experts have to justify deviation of their opinion from the mean. Each expert is providing his justification to all his colleagues. Then another round of weighting is carried out. The process continues till convergence of the weights is achieved. The Delphi method has been applied to different priority setting exercises (see for example von Oppen et al., 1996). The Delphi method, however, may only be appropriate for prioritizing overall criteria for research evaluation. Further application to research results on sublevels would be too complex and time consuming. No example of its application to agricultural research evaluation is known to the author.

1.2 Ex-post and Ex-ante Research Evaluation

In the previous section methods for identifying weights and priorities in agricultural research were distinguished by disciplines/methodological similarity. Another appropriate way of distinguishing wighting methods is by focussing on the context in which they are applied. It can be ex-post or ex ante.

- Ex-post evaluation involves in an evaluation of past research operations and mostly is related to outputs and outcomes which are known. In ex-ante research evaluation there is also more uncertainty. The objectives of ex-post analysis usually are (1) to improve management of on-going research activities, (2) to learn from past experiences in order to improve future planning (3) to reorient research to save time and money, (4) to satisfy the accountability requirements of funding agencies, and (5) to demonstrate research impact in order to ensure continued support for research from donors and government.

- Ex-ante evaluation looks at future research operations and the potential benefits they may induce. It is often used in research priority setting. Results of ex-ante research evaluation may come from simulations and they are often interpreted with the use of sensitivity analyses. In contrary to ex-post evaluation objectives of ex ante evaluation are (1) to identify research areas vital to agricultural development efforts, (2) to solicit government and donor support for vital research areas, and (3) to guide future research budgeting and resource allocation decisions.

Some methods may be more appropriate for ex-post and some for ex-ante evaluation, e.g. multiple criteria methods are hardly applied in ex-post; econometric analysis may be rare in ex-ante evaluation. In general, methods with simulation capacity tend to be more ex-ante than those without.

Table 1 shows the different research methods as they are used in ex-post or ex-ante research evaluation.

Table 1: Research evaluation methods in ex-post and ex-ante analysis

	Ex-post research evaluation	Ex-ante research evaluation
Econometric methods	✘✘✘	✘
Budgeting methods		✘✘✘
Economic surplus methods	✘✘✘	✘
Programming methods		✘✘✘
Multiple criteria methods	✘✘✘	✘✘✘

✘✘✘= frequently used
✘ = rarely used

1.3 The Soft Evaluation Approach

Research activities form part of a research system. It has been found, that those research systems should not be portrayed as hard blocks of institutionalized research operations which produce technologies as output (the "technology factory"). Research systems are rather soft and fuzzy spaces of competing and synergetic interaction involving a huge range of actors (e.g. government research organizations, universities, extension services, non-government organizations, private for-profit companies, farmers and farmers organizations) which in a partly structured and partly chaotic process generate new knowledge (Hartwich and Meijerink, 1999). In this regard a technology developed from agricultural research is just an indicator that new knowledge has been achieved.

The economically based methods have painted the evaluation problems within the context of the "technology factory". This led to a range of obstacles:

- The focus is on immediate technology impacts and not on the gain in substantial knowledge and expertise with limited technological impacts. In relation to applied research strategic research is discriminated.

- The gain in knowledge is not properly evaluated. The wide and complex system of research activities which produces a multiple set of research outputs is not taken in full consideration.

- The activities of respective research units are not properly evaluated because the focus is on overall technological impact and not on a comparison of different research units.

To overcome the above obstacles, any evaluation has to find ways to account for multiple outputs. In this regard some sort of weighting has to be applied in order to aggregate different outputs to an overall indicator of research performance. If primary data is not available this sometimes leads to the need to include subjective judgements. A method which is able to derive such weights allowing for reduced degrees of subjectiveness is the Analytic Hierarchy Process which is introduced in the following as an ex-post evaluation tool.

2. The Analytic Hierarchy Process – Background and Theoretical Aspects

From the above review it becomes clear that there is a rather limited choice of methods for research evaluation when the aim is to account for multiple input multiple output relations over a larger size of units of analysis. Multiple criteria methods and DEA seem to provide appropriate information on weights and priorities in agricultural research at reasonable costs. In the following we will focus on the use of the AHP method in ex-post research evaluation. It is a particular promising method as it can attribute weights to a set of very different research results. In the past, AHP has been particularly developed for decision support in management and the ex-ante dimension was the main focus. Newer research has applied AHP in ex-ante priority setting in agricultural (Braunschweig and Janssen, 1999). The main difference to ex-ante AHP is that the weights are not based on hypothetical views on future potential results of research but on the perception of relevance of research results derived in the past.

The AHP is a comprehensive, logical and structural framework which allows understanding of complex weighting problems by decomposing the problem in a hierarchical structure. In general AHP models can compare quantitative as well as qualitative information. Mostly the application of the AHP recognizes the knowledge of experts by making use for their subjective judgments. In this regard subjective judgments are transformed into quantitative data.

The AHP is of process character and can be devided into a sequence of different steps (Saaty 1990, Huizingh 1993). In Ex-post research evaluation the following 6 main steps have to be distinguished:

1. Defining the evaluation problem,
2. Defining and selecting the units of evaluation,
3. Identifying a set of alternative research outputs,
4. Identifying a set of relevant criteria,
5. Developing the hierarchical structure,
6. Collecting information and eliciting local and global priorities,
7. Preparing recommendations for action.

The mathematical algorithm behind the AHP is based on the Eigenvektormethod (Saaty 1980). This method is leading to the pairwise comparison procedure (step 5). Each pair of alternatives is compared with regard to the respective next level criteria. Those comparisons can be verbalized. The question for example would be: With regard to criteria A, how much more important is alternative a over alternative b (or vice versa), or, transforming it to an example from agricultural research priority setting, with regard to the criteria of food security how much more important is research on problem a over research on problem b. A ratio scale between 1 and 9 is used (see Table 2) with a value of 1 indicating indifference and a value of 9 indicating very strong preference. This scale has been derived by Saaty from insights of psychological science and has been defended on empirical grounds (Harker, 1998). It has been shown that the human mind uses multiplicative dimensions when comparing two alternatives. If, for example, somebody feels that alternative a is much more important than alternative b, this would mean that a is five times more important than b. This scale can be applied to the dimension of importance as well as to likeliness and preferredness. The term importance, here, refers to the state of an alternative of being important or significant or lead to consequences, likelihood refers to the state of an alternative to be likely or probable, i.e. the chance that an alternative occurs, and preferredness refers to the condition or state of an alternative of being preferred.

Table 2: Scale of Measurement for the AHP

Importance	Likelihood	Preferredness	Numerical Value
Equally important	Equally likely	Equally preferred	**1**
Moderately more important	Moderately more likely	Moderately more preferred	**3**
Strongly more important	Strongly more likely	Strongly more preferred	**5**
Very strongly more important	Very strongly more likely	Very strongly more preferred	**7**
Extremely more important	Extremely more likely	Extremely more preferred	**9**
Intermediate values	Intermediate values	Intermediate values	**2,4,6,8**

After deriving the judgements from interviewing individuals or groups AHP involves in entering the judgments into a matrix. A simple n=3 matrix (where n is the number of alternatives) using some hypothetical values is shown in Table 3. The values in the cells of the matrix represent the weighting results (it reads as the relation between the alternative in the row to the alternative in the column). The cell of line 1/column 2 in Table 3, for example, shows that alternative a is judged to be between equally and moderately more important than alternative b. The values in the diagonal of the matrix are always one (when compared with itself, each alternative has equal importance) and the lower triangle values of the matrices are the reciprocal values of the upper triangle. Therefore, pairwise comparison is required for only half of the matrix, excluding the diagonal values. The number of comparisons to undertake (with regard to one criteria) can be computed from the formula [n (n-1)], where n is the number of alternatives.

Table 3: Matrix of pairwise comparisons

	Alternative a	*Alternative b*	*Alternative c*
Alternative a	***1***	***2***	***1/3***
Alternative b	***1/2***	***1***	***1/4***
Alternative c	***3***	***4***	***1***

Inconsistency: 0,00

If one compares three alternatives a, b and c he may consider three types of comparisons: a/b, a/c, and b/c. In this case the third comparison b/c is transitive, i.e. it is not necessary because the two other comparisons sufficiently establish the triangular relation between the alternatives. However, there are problems in the transitivity of human judgements. For example, if alternative a is judged more important than b, and b is judged to be more important than c, then a should also be judged better than c. In reality humans may judge c being more important than a. The result of this inconsistency are perverted scores of importance.

In the case of three alternatives, AHP would always exercise not only two comparisons but also the third in order to test the degree of inconsistency of the judgment. For measuring the inconsistency of judgements Saaty developed a consistency index, which compares the inconsistency of the eigenvectors calculated from the matrix with calculated inconsistencies of a set of randomly chosen matrices of the same size (Haedrich, 1986). The thus derived value of inconsistency is zero if the interviewees judgement is totally consistent and is one if the judgments are completely inconsistent. It has been shown, that humans usually judge with a certain degree of inconsistency. It is therefore not the aim of the AHP to come to 0,00 consistent judgments but to allow some degree of inconsistency (up to 0,10) to monitor reality (Saaty, 1980). In this paper, inconsistency ratios are given under the respective Tables.

In AHP nomenclature weights derived from the comparison process are called priorities. Local priorities are derived from comparisons with respect to one particular criteria. Global priorities are derived from aggregating all local priorities, they constitute the final results of the AHP.

3. An Example from Agricultural Research Evaluation

In the following we illustrate the AHP procedure for the background of an agricultural research evaluation problem. In country X agricultural research has been carried out for the last 30 years by different national and international research centers and, to a lesser extent, by the national university. we assume, that one particular field of research which is considered to be of primary importance to food supply and rural welfare in the country X is research on sorghum. Three national research stations, an international research station (sponsored by the World Bank, FAO and the EU) and a field station of the university have involved in sorghum research. However, results are not satisfying the needs of the farmers and adoption rates on small scale farms (building the majority of the sorghum producers) are low. National research managers think about restructuring sorghum research in the country. For this purpose, they need information on past research performance of the different research stations involved in agricultural research and their respective comparative advantages in sorghum research. Information on the number of research outputs (as for example varieties developed, agronomic techniques developed, master and Ph.D. thesis conducted, scientific papers written) are fully available. However, the question is how to judge those outputs. How to compare apples with oranges? What would be the appropriate weights to those outputs? What are the respective contributions of sorghum research to attain the overall goals of sorghum research? The AHP procedure is able to give answer to those questions. In the application of the AHP to our problem we will follow the 6 steps mentioned above.

Step 1: Defining the Evaluation Problem

In country X resources for research are extremely scarce. All institutes involved in sorghum research form the joint national sorghum research force. Though sorghum is one of the major crops as well as one of the major staples in the country it is not easy to decide how much resources should be attributed to sorghum research. Against the interests of big commercial cotton farmers, exporting companies and certain groups in the government profiting from cotton export, small scale farmers, farmers organizations, NGOs, and extension service organizations have convinced the government (the Ministry of Agriculture) and international donors, that sorghum research is of primary importance and certain resources are allocated to it (the priority of sorghum research in relation to other research is not dealt with here).

The overall goal of sorghum research is that through improved sorghum based cropping systems welfare in the country is improved. However, there are many alternatives in sorghum research. Plant material could be improved through adaptive breeding of varieties from other countries, marketing of products may have a huge potential of improvement, research on plant hygiene may alleviate stress from diseases and parasites, and research on agronomic techniques like spacing and inter-cropping may be highly important to increase yields. We thus define our evaluation problem as to analyze all past research activities in sorghum research with regard to a set of relevant criteria.

Step 2: Defining and Selecting the Units of Evaluation

An inventory reveals that three research organizations, two universities and one government research organization, with altogether 10 stations, 2 university departments and 8 government research stations, and a number of 100 researchers were involved in sorghum research. Evaluation can take place on the system, organization, station, program, project, activity and individual researcher level. The participants in sorghum research on those levels could be defined as units of evaluation. However, the aim in the study was nor on personal researcher performance evaluation neither on a comparison of organizations and stations but on project unit in which interaction between institutions is possible and where synergies can be accounted for. Therefore the project was chosen as unit of evaluation.

Step 3: Identifying a Set of Alternative Research Outputs

The question here is in what areas sorghum research has been carried out in the past. An inventory reveals that in the last 30 years there were at least 60 projects with multiple objectives which involved in one way or another in sorghum research. However, how to group them under some conclusive topics. For this purpose an identification workshop has been carried out attended by researchers, practitioners, farmers, representatives of NGOs and farmers organizations, and extensionists. Involving metaplan techniques, the following main topics of sorghum research were identified: (1) breeding, (2) agronomy, (3) plant hygiene, (4) post harvest. Of course, more topics and sub-topics could easily be defined. For illustrative purposes we limit the number of alternatives, i.e. research activities to 4. It is important to note that in AHP the defined alternatives (and criteria) should not be interdependent or overlapping. In our example there could be overlapping between breeding and post-harvest research as breeding can be focussed also on the improvement of post-harvest qualities of sorghum. To avoid overlapping, for example, post-harvest research could be defined as all research on post harvest technologies which does not involve in breeding.

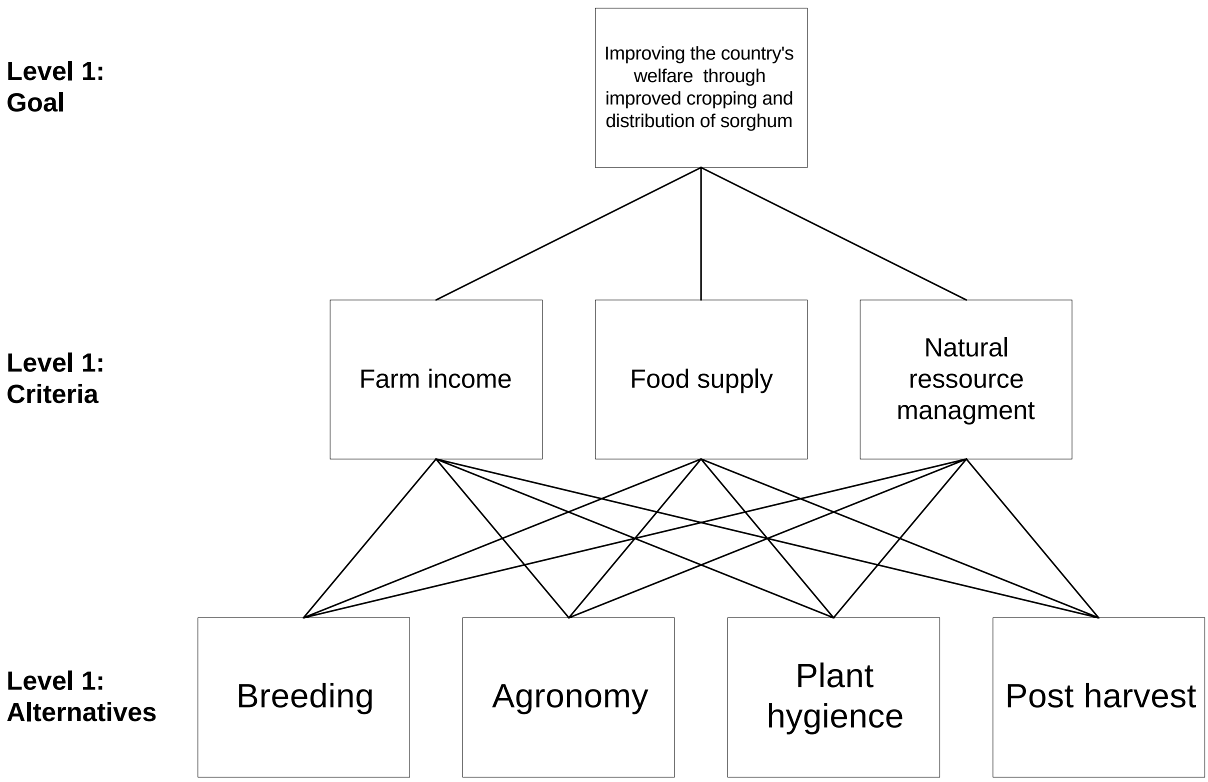
Step 4: Identifying a Set of Relevant Criteria

The purpose of sorghum research has to be seen in the potential of sorghum to contribute to a better food supply of the people and contribute to farm income increases. Furthermore, the sustainability of agricultural production for the background of deteriorating natural resources is of growing importance. Along this arguments, three criteria were defined by the participants in the identification workshop on which the above alternative of sorghum research activities were supposed to be judged: (a) contribution to farm income increases, (b) contribution to improved food supply, and (c) contribution to natural resource management. Again, more criteria could be defined, but for keeping the example simple the number of criteria is reduced to those three.

Step 5: Developing the Hierarchical Structure

In this step the decision problem is decomposed into a logic hierarchical structure. Developing hierarchies is not always easy. It usually needs some experience. Problems occur in the reflection which alternatives and criteria are relevant and to which levels of the hierarchy they belong. This is particularly difficult if the hierarchy involves many levels. Often, redefinition of alternatives and criteria (step 2 and 3) have to be undertaken. In our example we come to a simple three level hierarchy involving only goal (level 1), criteria (level2), and alternatives (level 3).

Figure 1: The Hierarchical Structure of the Decision Problem



Step 6: Collecting Information and Eliciting Local and Global Priorities

Once a structure that models the decision problem has been established, the next stage is to evaluate the alternatives as well as the criteria by making pairwise comparisons. The pairwise comparison procedure on level 1 has been carried out in a group discussion with each a farmer, extension and research representative. This was done in a special meeting after the identification workshop. The workshop revealed one consistent group judgement result. The judgments could have been alternatively derived from personal interviews with the representatives. When several individuals provide judgements one may aggregate individual judgments. Forman and Peniwati (1996) found that when pooling judgements one should take the geometric average of individual judgments as it satisfies the reciprocity requirement implying a synergistic aggregation of individual preferences in such a way that the group becomes an individual.

Table 4: Matrix of pairwise comparisons of criteria

	Farm income	*Food supply*	*NRM*
Farm income	1	2	8
Food supply	1/2	1	6
NRM	1/8	1/6	1
Total	13/8	19/6	15

Table 4 illustrates judgements of the interviewees in a matrix form. The cell in the first row and the second column indicates, that asked to the question of what criteria was more important in sorghum research, "farm income" or "food supply", the interviewees judged "farm income" being between equally and moderately more important (first line / second column in Table 4) (score 2 according to Table 2). The cell of the first column and the second line indicates the reciprocal value of that relation, i.e. "farm income" being between equally and moderately less important than "food supply" (score ½). "Farm income" was judged being between very strongly and extremely more important than "natural resource management" (score 8) and "food supply" strongly and very strongly more important than "natural resource management" (score 6).

Table 4 is than transformed to the normalized matrix (Table 5)by division through the column total.

Table 5: Normalized matrix of pairwise comparisons of criteria

	Farm income	*Food supply*	*NRM*
Farm income	8/13	12/19	8/15
Food supply	4/13	6/19	6/15
NRM	1/13	1/19	1/15
Total	1	1	1

From Table 5 weights (local priorities) for the alternatives can be derived from simply calculating the arithmetic mean (see Table 6).

Table 6: Computation of local priorities of criteria

		Local priorities
Farm income	[(8/13)+(12/19)+(8/15)]/3	= 0,593
Food supply	[(4/13)+(6/19)+(6/15)]/3	= 0,341
NRM	[(1/13)+(1/19)+(1/15)]/3	= 0,066
Total		1,000

Inconsistency: 0,02

After prioritizing the criteria the results from the group discussions on the alternative research outputs are derived. The questions are asked in the way: With respect to farm income which research outputs have been more important, those in "agronomy" or those in "breeding" and to what degree were they important. The judgements revealed, for example, that results in "agronomy" were regarded to be much more important than results in "breeding". The set of pairwise comparison judgements are shown in the following Tables 7 to 9. For simplification there are no extra tables illustrating the computation of the local priorities but they are added to the tables in an extra column.

Table 7: Matrix of pairwise comparisons of alternatives with respect to farm income

	Breeding	Agronomy	Plant hygiene	Post harvest	Local priorities
Breeding	1	5	4	2	0,483
Agronomy	-	1	1/3	1/5	0,066
Plant hygiene	-	-	1	1/3	0,136
Post harvest	-	-	-	1	0,315

Inconsistency: 0,04

Table 8: Matrix of pairwise comparisons of alternatives with respect to food supply

	Breeding	Agronomy	Plant hygiene	Post harvest	Local priorities
Breeding	1	1	1/5	1/6	0,079
Agronomy	-	1	1/2	1/4	0,107
Plant hygiene	-	-	1	1/4	0,236
Post harvest	-	-	-	1	0,578

Inconsistency: 0,07

Table 9: Matrix of pairwise comparisons of alternatives with respect to NRM

	Breeding	Agronomy	Plant hygiene	Post harvest	Local priorities
Breeding	1	1/5	1/3	2	0,123
Agronomy	-	1	2	5	0,519
Plant hygiene	-	-	1	2	0,263
Post harvest	-	-	-	1	0,095

Inconsistency: 0,04

From Table 7 to 9 we see, that with respect to farm income, food supply, and natural resource management the most important research topics are breeding (48%), post harvest (58%) and agronomy (52%) respectively. However, we still do not know which research topic is the overall most important. In a last stage of step 5 we need to apply the weights of the criteria from Table 6 to come to the overall priorities of the alternative research topics, in other words local priorities are transformed into global priorities. This process consists in multiplying the local priorities with the weight of each criterion and then sum up the products to yield the final ranking of the alternatives (see Table 10).

Table 10: Computation of the global priorities

	Farm income	Food supply	NRM	Global priorities
	(0,593)	(0,341)	(0,066)	
Breeding	0,483	0,079	0,123	(0,593x0,483)+(0,341x0,079)+(0,066x0,123) = 0,322
Agronomy	0,066	0,107	0,519	(0,593x0,066)+(0,341x0,107)+(0,066x0,519) = 0,109
Plant hygiene	0,136	0,236	0,263	(0,593x0,136)+(0,341x0,236)+(0,066x0,263) = 0,178
Post harvest	0,315	0,578	0,095	(0,593x0,315)+(0,341x0,578)+(0,066x0,095) = 0,390

Overall Inconsistency: 0,04

We can see that according to the judgments of the experts involved the most important research activity related to sorghum is research on post harvest technologies (39%). This is followed by breeding (32%), plant hygiene (18%), and agronomy (19%).

The results from Table 10 can be used for example to reflect on the allocation of future funds. If post harvest has a priority of 39% then in a static way also 39% of the funds for sorghum research could be attributed to post harvest research activities. At least, it may be reasonable that the rankorder of the size of funding for each topic is be in line with the rankorder of the priorities. However, other criteria as importance to a region, future development potential, equity contribution may have to be included in the decision of policy makers on future funding of research.

Further interpretation of the results can be carried out if information on the present research funding situation would be available. In our example the total funds for sorghum research in the country amount to 1.000.000 USD and are distributed according to Table 11. In this case we can postulate that current research priorities do not reflect in the attribution of funds. The research manager who consulted the AHP analysis concludes that a reallocation of funds away from breeding and agronomy towards post harvest research would be useful.

Table 11: Example of funding and priorities in sorghum research

	Distribution of funding in Sorghum research activities (in %)	***Priorities (in %)***
Total	4.000.000 USD	100%
Breeding	55%	32%
Agronomy	20%	11%
Plant hygiene	20%	18%
Post harvest	5%	39%

If overall performance of research stations has to be evaluated the number of results in a research category has to be rated against the derived global priorities. Practical evaluation procedures would have to distinguish between different research results within the categories. In breeding, for example, results as new varieties, adapted varieties, or papers published may have to be distinguished. Experts would have to weight, for example, the importance between new varieties developed and existing varieties adapted to local conditions.

In Table 12 an example of research results as they have been achieved in the last 10 years (already agglomerated and weighted) by 5 different research stations involved in sorghum research in country X is given. With this information on the overall performance of different research stations in sorghum research can be deducted. Of course, the weighted outputs of Table 12 would have to be related to the resources used (human, financial and physical) in order to see, if the stations are deriving their results efficiently. However, Table 12 gives some very useful tendencies on research performance. The performance indicator for example could be used by the research manager to channel future funds according to comparative advantages. E.G. Station 3 and 5 may get more funds because they are good in generating post harvest research results.

Table 12: Example of funding and priorities in sorghum research

	Station 1 Government	***Station 2 Government***	***Station 3 Government***	***Station 4 International***	***Station 5 University***	***Priorities (in %)***
	No of research results (technologies)					
Breeding	4	0	0	6	0	32%
Agronomy	3	3	4	2	5	11%
Plant hygiene	2	0	0	4	0	18%
Post harvest	0	0	4	1	6	39%
Weighted output	1,97	0,33	2	3,25	2,89	100%

Step 7: Preparing Recommendations for Action

Results like above may be questioned by the different stakeholders involved. It is therefore important to transform the results into recommendations which are robust. Only then policy makers will be using the results from the AHP in future budgeting and resource allocation decisions. In general, the recommendations should be formulated in line with the objectives of the research evaluation (see 1.2 for objectives of ex-post research evaluation). In the example, one may assume that plant breeders will criticize the results of the evaluation with the argument that “Only through new varieties development of the sorghum sector can be achieved.” They would argue that “New varieties will solve problems in agronomy and plant hygiene as well. Markets for the improved sorghum products are abundant and no research on post harvest technologies needs to be undertaken if only the new varieties were applied.” It is therefore important to define clearly who will participate in the judgment process and with what weight. The research manager, in our example, will be able to say, that a clearly defined group of stakeholders came up with the results that post harvest research (not involving in breeding) is more important than breeding.

In order to see how robust the results are a sensitivity analysis can be undertaken providing information under which circumstances post harvest research is becoming less important. For this purpose the weights of the criteria (farm income, food supply, and NRM) are changed. This simulation can easily be carried out with the use of the sensitivity function of the Expert Choice Software. It is however difficult to visualize this method in a brief paper. The results of the sensitivity analysis reveal that only if the priorities for farm income would raise to 72% breeding will become more important than post harvest research. Alternatively raising the priority for NRM to 40% would make agronomy research more important than breeding raising it to 45% which is even more important than post harvest research.

4. Discussion

AHP uses mathematical algorithms to transform qualitative subjective judgements into quantitative data (weights). That particular characteristic makes it an interesting alternative to quantitative techniques, particularly in the field of agricultural research where complex decision structures related to multiple criteria exist.

4.1 Strength of the Method

As a method of identification of weights and priorities AHP provides a range of advantages. It is evident, that the stakeholders involved come to rational results due to the logic decomposition of the problem which suppresses personal preferences. A focus on goals is achieved through the use of structured hierarchies. Another advantage is that subjective qualitative information is transformed into quantitative data which can be used for management decisions. Through this, a de-emotionalization of the priority setting process can be achieved, particularly when judgments are derived in groups. Prioritizing among alternatives in a structured setting is achieved although there is lack of sufficient quantitative data. The results, a small set of percentage numbers, are rapidly computed and easy to interpret. In addition, an important feature of AHP is, that the costs compared with other methods (e.g. econometrics, cost-benefit, programming) are low.

Econometric regression analysis derives estimated elasticities of variables which can be used as indicators for importance of those variables. If sufficient data is available alternatives can be compared due to those elasticities. AHP may come to very different results as regression analysis and therefore enrich understanding of the problem from a less quantitative point of view. Hassell et al. (1992) used AHP as an alternative to regression analysis to assess the relative importance of performance evaluation criteria in an accounting context. They found, that the weights assigned to the criteria and the criteria rankings differ from those in prior regression–based studies suggesting that the initial regression-based studies may not have provided definitive conclusions about the relative importance of performance evaluation criteria.

Another advantage is, that participants of the evaluation process gain insights into the critical areas of the decision process. Karbhari (1994) found that AHP is a viable means for rapid decision-making in a team of very different scientists, in his case engineers in the car industry. He found, that AHP provides results consistent with those derived though more in-depth and time-consuming approaches as it forces structured decision-making and allows the team to evaluate biases and trends within groups. Several extensions of the AHP exist. One promising approach has been suggested by Rosenbloom (1996). He involved in a probabilistic approach treating the input data in the AHP as random variables. This approach provides a means to test the statistical significance of the final score and has the advantage of allowing a probabilistic interpretation of the final scores.

4.2 Limitations of the Method

The AHP process has been widely discussed in the literature by scientists and practitioners (see for example Karbhari, 1994 or Brinkmeyer and Müller, 1994). Though mathematically proofed there is still debate on the possibilities of application and the degree to which the results are relevant. Most often critique comes from the argument that very different alternatives cannot be compared. AHP is capable of any comparison as long as the target group is able to carry out a further decomposition of the problem (and the criteria) in an analytic hierarchy. There may be, however, very complex problems resulting in a very complex process and the amount of feasible pairwise comparisons to be conducted in a certain time is a limit.

Some scholars have criticizing that AHP does not give any constructive guidance to the structuring of the problem, for example the research evaluation problem. The question, for example, if the structure of a hierarchy needs is flat (with few levels and many elements on the levels) or tapered (with many levels and few elements on one level) is of crucial importance to the result of the weighting process. Empirical work indicates that different methods of designing the hierarchies of the same problem may lead to contrasting results (Adelmann et al., 1986; Stillwell et al. 1987). In response to the critique it has to be noted that its authors assume the existence of an objectively perceivable problem. However, it is more likely that problems are perceived subjectively by individuals. The structure of the hierarchy than has to follow the perception of the individual (or the group of individuals) and there is no possible alternative to this problem formulation (Brinkmeyer and Müller, 1994). The design of another hierarchy presupposes the existence of another problem. Also, AHP may be the tool to arrive at a consensus on the structure of a problem, and thus alllows to overcome the initial diversity in perspectives.

Another issue relates to the way of aggregating opinions. Either few or many individuals can participate in the AHP. The aggregation of weights can be achieved either compensatory or non-compensatory among the group carrying out the performance rating (Arrington et al. 1982). The question is when to merge those ratings. They can be either merged within a group discussion; what will be fit in the mathematical algorithm is already an aggregate of the peers' judgements. Alternatively, every individual can carry out his rating separately. AHP then calculates the local and global priorities for each individual. Only at the final stage the priorities will be aggregated with the use of geometric means. More complexity is achieved when different peers or groups carry out their ratings on different levels, e.g. keyactors in management judge the importance of criteria, and researchers judge the importance of alternatives. What approach to take is largely due to the type, purpose and scope of the evaluation study. If very different stakeholders, e.g. farmers and researchers, are participating and it is expected that one group will suppress the opinions of another then separate rating is useful. Separate rating is also useful when the peers cannot be brought together easily due to geographic distance or time constraints. Group decisions are useful when arguments of different participants should be brought together to encourage the discussion. Evidence from different studies shows that the analytic hierarchies established by AHP are well suited for evoking fruitful discussions and participants are often supportive to the resulting weights even if they do not reflect fully their interests. Locket et al. (1986), for example, claim that the interactive character of the AHP used in group decisions encourages the development of subjective estimates and their subsequent discussion. They saw, that the exercise strengthened the group because individuals could see views of their colleagues and their reasons without necessarily agreeing with them. Anders and Müller (1995) regard AHP as useful approach to structuring communication in a diverse research group. For group discussions a special Software is available called Team Expert Choice.

Intensive critique and discussion has been articulated vis a vis the rank reversal problem. Rank reversal means that the rank of an alternative resulting from AHP may change when another alternative is added to the initial group of alternatives compared. Saaty (1987) responded to this critique saying, that with introduction of a new alternative also new information is included in the model. In this regard, the decision problem has to be rethought, and the resulting ranks of alternatives may change. However, scholars have identified solutions to cope with the problem in a methodological way. They include: Referenced AHP, normalization to the maximum entry, normalization to minimum entry, and linking pins. Explanations of these approaches can be found in Schoner et al. (1993). However, with regard to practical decision making those approaches tend to lack some practicability (Brinkmeyer and Müller, 1994). The latest software for AHP, Expert Choice © (Version 9,0) includes an alternative "ideal synthesizing mode" which allows that the sum of alternatives adds to more than one. In this respect it is not necessary to newly calculated priorities of existing alternatives when introducing a new alternative. Through this the rank reversal problem is excluded.

A limitation of AHP is that the procedure of pairwise comparisons is time consuming. [(n x (n-1) x j] pairwise comparisons have to be conducted per level with n = number of comparisons and j = number of criteria. Some interviewees may find it somewhat tedious to go through that amount of pairwise comparisons (Lockett et al, 1986). Also, experts involved begin to reflect too excessively on the problem. Then, the given answers are not intuitive as they should. Related to this critique is the argument that after a certain time, experts may loose interest in the interviews. To avoid this, some authors suggest simpler multiple criteria methods to be applied. It is an advantage that ranking derived from scoring procedures can be conveniently combined with AHP.

4.3 Applying AHP to Agricultural Research

AHP has been widely used in management and decision support. Just to name an example, Locket et al. (1986) have been using AHP in the selection of research portfolios in the pharmaceutical sector. It has been also used in research management and agricultural research priority setting. Anders and Müller (1995), for example, have been using AHP as a means of developing the agenda for a long-term cropping experiment. They elicited disciplinary specialists' preferences and measured for alternative research options in an interdisciplinary project at the International Crops Research Institute for the Semi Arid Tropics (ICRISAT) which was designed to produce a multi-technology package in India. A hierarchy of options was developed with soil types at the first level, cropping systems at the next and treatments at the lowest level. Preferences for experiment options were elicited in two rounds from 20 scientists using structured questionnaires. Data from each round were pooled to derive overall weights for the respective treatments. The method enabled to pinpoint on some possible causes of difficulties in earlier attempts to overcome paradigmatic differences.

In a University of Hohenheim / ISNAR project on research performance assessment in Cameroon and Tanzania, agricultural research performance of national research organizations and national universities has been analyzed. First steps involved the development of a cumulative performance indicator (CPI) to describe the performance of research institutions (Hartwich, 1998). This approach combined quantitative data on research outputs with weighted scores based on qualitative assessments of the research using AHP. Analysis of an institution's research performance yield a weighted outputs indicator, which can be divided by the institution's budget or another input measure to give the cumulative performance indicator (Alex, 1998). In a further step, a methodology was developed which combined Data Envelopment Analysis with AHP. Within categories, the AHP was used to weigh different outputs, then, to avoid further agglomeration of categories which cannot be compared, DEA was applied calculating Pareto efficiencies. This led to an overall weighted technical efficiency indicator with which the research units under investigation could be compared (Hartwich, forthcoming). The indicator proved to be a very informative and easy to interpret measure. Its values corresponded to observations in the field. It clearly identified areas of comparative advantages and weaknesses of research units and it guided to management decisions which could lead to higher performance of the units and the whole research system. It also identified types of research organizations (Universities, NAROs, international Centers) which were particularly technically efficient.

As was shown through the example and the discussion, the AHP is an interesting new tool when deriving of weights and priorities in agricultural research is required. Particular assets of AHP are:

- Agricultural research responds to multiple objectives. AHP derives scores in relation to even a large sets of objectives simultaneously.
- In an evaluation of multiple agricultural research operations traditional econometric and economic surplus approaches require a burdensome amount of data and time. Also these approaches are not able to distinguishing between the different research units involved in the research operation under investigation. Interpretation of the results and use in management decisions is therefore limited. AHP outperforms those methods on the account of using subjective perceptions by peers, experts, and stakeholders on the social benefit of research operations.
- AHP outperforms the results of often subjectively biased multiple criteria methods as it requires structured hierarchies of the evaluation problem which provide formalization and consistent judgements.
- The AHP is a process which involves a certain amount of peers, experts, and stakeholders. As many stakeholders are involved it avoids anonymous communication of the results of an evaluation process by an evaluation unit. Through this a higher degree of acceptance may be achieved. Further, when group decision procedures are applied consensus building among different stakeholders may be an asset.

A key characteristic of AHP is that it relies on existing alternatives. This applies to AHP as to other priority setting methods, particularly the more they relate to data. Existing alternatives are stronger in the mind of people than alternatives which only materialize in the future and which are difficult to describe. The problem with this is that in research crucial problems may lie in new fields that were not researched at all or that did not produce any research results for the time being. In this regard the AHP is providing biased conservative weights. This is particularly a shortcoming of AHP in ex-ante priority setting (Braunschweig, 1998). Some creativity approaches may be used to generate alternatives, but once the alternatives are processed in a structured manner, they always have problems maintaining their prominence for the future. However the disadvantage of conservative weights is not evident in ex-post analysis. It is important to note in this circumstance, that AHP should not be seen as a tool for isolated, one-time application, but rather as a process that has ongoing validity and usefulness to an organization (Saaty, 1989).

One problem particularly related to research evaluation using AHP is, that it may be difficult to evaluate research fields which overlap. E.g. sorghum agronomy may overlap with soil nutrition research. In this case it is important to refer to the overall goal of the hierarchy. The question is if sorghum research or soil nutrition research is evaluated. Comparison of overlapping research is not possible. Even comparison between non overlapping alternatives e.g. Sorghum and Maize research is possible if they relate to one overall goal (for example providing food and income to the rural sector).

AHP weights can also be used to derive funding decisions in agricultural research. Grizzle (1987), for example, examined the suitability of the AHP in performance monitoring systems at the state and local levels of government school teachers in Florida. He found the AHP to be a suitable technique for eliciting weights for developing overall performance scores for public agencies and even advocates the usefulness of tying funding levels to performance levels. However, in funding decisions, there are some practical considerations which may restrict the use and the relevance of AHP evaluations. For example, national research policy makers may still depend on donor priorities as much of the resource allocation in agricultural research tends to be donor driven. It is to see if in the future donors will be able to subordinate their research priorities to priorities derived from a national priority setting process for the whole country.

Conclusion

In the past econometric and economic surplus methods have been used in research evaluation, with mixed applicability of their results. They have been accepted, though only in parts, in research evaluation. The results of those approaches are powerful tools in overall impact assessment of agricultural research while justifying investments in research. However, they are not able to derive information on research performance of specific research units. They are also not able to account for multiple research outputs as they are derived in disperse research systems. AHP is capable of doing both. Overall, the AHP presents a powerful tool for weighting and prioritizing alternative research outputs. It is adapted to structuring of complex weighting problems, particularly when a set of discrete alternatives is related to a multiple and complex set of objectives, as being the case in agricultural research. Experience in ex-ante agricultural research evaluation are available and provide valuable results. The discussion shows that the method is also very useful in ex-post research evaluation because judgements on past research result, as derived from the AHP inherent measurement and decomposition process, may be quite accurate.

Different problems exist in the application of the method as for example the problem of designing consistent hierarchies, the large amount of pairwise comparisons to conduct, the rank reversal problem, and the issue of aggregating opinions of different peers. Those problems are not ruling out the advantages of AHP in general and suitable solutions to the problems do exist. The limitations may rather lie in the fact that researchers and administrators remain critical when being evaluated, regardless to whatever set of objectives. The argument from this side is, that research results may respond to objectives which only become visible in the future after they are derived. However, in times of constraint resources national agricultural research systems become highly under pressure. Particularly in developing countries, where science focuses on the adaptive side of the research spectrum, research not related to any objectives may be considered as a waste of resources. This opinion is likely to be adopted more and more by national governments and international donors. Evaluation schemes using AHP would constitute a useful tool to guide research towards objectives. Part of the limitations of AHP may lie in the fact, that many scholars are still not familiar with the methodology. Further application and testing of AHP in ex-post agricultural research evaluation would therefore be useful.

Literature

Adelmann, L.; Sticha, P.J.; Donnell, M.L. (1986). An Experimental Investigation of the Relative Effectiveness of Multiattribute Weighting Techniques. In: Organizational Behavior and Human Performance, Vol. 33, pp 243-262.

Anders, M.M.; Müller, R.A.E. (1995). Managing Communication and Research Task Perceptions in Interdisciplinary Crops Research. In: Quarterly Journal of International Agriculture, Vol. 34, p. 53-69

Arrington, C.E.; Jensen, R.E.; Tokutani, M. (1982). Scaling of Corporate Multivariate Performance Criteria: Subjective Composition Versus the Analytic Hierarchy Process. In: Journal of Accounting and Public Policy, Vol.1, pp. 95-123.

Braunschweig, T. (1997). An illustrative Example for research Priority Setting Using the AHP Approach. ISNAR, The Hague, Unpublished

Braunschweig, T.; Janssen, W. (1998); Establecimiento de Prioridades en la Investigacion Biotechnologica mediante el Proceso Jerarquico Analitico. ISNAR, Research Report 14. International Service for National Agricultural Research, The Haag, The Netherlands.

Brinkmeyer, D.; Müller, R.A.E. (1994). Entscheidungsünterstützung mit dem AHP. Zeitschrift für Agrarinformatik 5/94.

Falconi, C.A. (1993). Economic Evaluation. In: Horton, D. et al.; Monitoring and Evaluating Agricultural Research: A Sourcebook. ISNAR & CABI, University Press, Cambridge, UK, 1993.

Forman, E.; Peniwati, K. (1996). Aggregating Individual Judgments and Priorities with the Analytic Hierarchy Process. ISAHP, Vancouver, Canada, July 12-15, 1996.

Grizzle, G.A.(1987). Pay for Performance: Can the Analytic Hierarchy Process hasten the Day in the Public Sector? In: Mathematical Modelling 9(3-5):245-250.

Haedrich, G.; Kuß, A.; Keilkamp, E. (1986). Der Analytic Hierarchy Process. In: WiST, Vol 3., pp 120-126.

Harker, P.T. (1998). The Art and Science of Decision Making: The Analytic Hierachy Process. In: Golden, B., Wasil, E.; Harker, P. (eds.). The Analytic Hierarchy Process: Applications and Studies. Springer, New York, U.S.

Hartwich, F. and Meijering, G. (1999). New Views on the Generation of Agricultural Knowledge: Revising the systems approach to national agricultural research. Forthcoming: ISNAR Discussion Paper, ISNAR, The Hague, The Netherlands.

Hartwich, F. (1998). A Cumulative Indicator for Measuring Agricultural Research Performance: Accumulating performance measures of agricultural R&D operations in a developing country. ISNAR Discussion Paper, No. 98-1, ISNAR, The Hague, The Netherlands.

Hartwich, F. (forthcoming). Performance Evaluation in Agricultural Research: Results from Low-input Low-output Research Systems in Sub-Saharan Africa.

Horton, D. et al. (1993). Monitoring and Evaluating Agricultural Research: A Sourcebook. ISNAR & CABI, University Press, Cambridge, UK.

Huizingh, K.R.E. (1993). Analytic Hierarchy Process: Een methode om appels met peren te vergelijken. Tijdschrift voor Marketing, March 1993.

Karbhari, V.M. (1994). The Analytic Hierarchy Process: A viable decision tool for composite materials? International Journal of Technology Management, Vol. 9, No 1.

Krawiec, F. (1984); Evaluating and Selecting Research Projects by Scoring, In: Research Management, No 2.

Locket, G.; Hetherington, B.; Yallup, P.; Stratford, M.; Cox, B. (1986). Modelling a Research Portfolio Using AHP: A group decision process. In: research Management, Vol. 16, No. 2.

Mills, B.; Kamau, M. (1998). Methods for Prioritizing Research Options. In: Mills, B. (Ed.) Agricultural Research Priority Setting: Information Investments for Improved Use of Resources. International Service for National Agricultural Research, The Hague, The Netherlands.

Norton, G.; (1993); Scoring Methods. In: Horton, D. et al. (eds). Monitoring and Evaluating Agricultural Research: A Sourcebook. ISNAR & CABI, University Press, Cambridge, UK.

Saaty, T.L. (1980). The Analytic Hierarchy Process. Mc Graw Hill, U.S.

Saaty, T.L. (1989). Group Decision Making and the AHP. In: Golan, B.L.; Wasil, E.A.; Harker, P.T. (eds.). The Analytic Hierarchy Process: Applications and Studies. Springer Verlag, Germany.

Saaty, T.L. (1990). How to make a decision: The Analytic Hierarchy Process. In: European Journal of Operational Research, Vol. 48.

Schoner, B.; Wedley, W.; Choo, W.C. (1993). A unified approach to AHP with linking pins. In: European Journal of Operational Research, Vol. 64, pp. 384-392.

Stillwell, W.G.; Winterfeld, D.; John, R.S. (1987). Comparing Hierarchical and Nonhierarchical Weighting Methods for Eliciting Multiattribute Value Models. In: Management Science, Vol 33, pp 442-450.

Von Oppen, M.; Heidhues, F.; van den Akker, E. (1996): Macro-Economic Evaluation of Technical and Institutional Innovations: Structural Adjustment, Food Security and Infrastructure in the Economic Region of Niger, Benin and Border Countries. Special Research Programme 308, Report of Results 1994 - 1996, pp.501-530, University of Hohenheim, Stuttgart.

Recent Discussion Papers

No. 01/99	F. Heidhues and G. Schrieder, *Rural financial market development.*
No. 02/99	F. Heidhues, C. Karege, B. Schaefer and G. Schrieder, *The social dimension of policy reforms.*
No. 03/99	F.Heidhues, W. Erhardt, A. Gronski and G. Schrieder, *The social dimension of reforms and World Bank case studies.*
No. 04/99	W. Erhardt*, Credit for poor and low-income entrepreneurs in urban and rural Northern Thailand.*
No. 05/99	J. Senahoun, F. Heidhues and D. Deybe, *Structural adjustment programs and soil erosion: a bio-economic modelling approach for Northern Benin.*
No. 06/99	Thanda Kyi and M. von Oppen, *An Economic Analysis of Technical Efficiency of Rice farmers at Delta region in Myanmar*
No. 07/99	Schrieder, G., Munz, J., and R. Jehle, *Rural regional development in transition economies: Country case Romania*
No. 08/99	Hartwich, F., and T. Kyi, *Measuring Efficiency in Agricultural Research: Strength and Limitations of Data Envelopment Analysis*
No. 09/99	Hartwich, F., *Weighting of Agricultural Research Results: Strength and Limitations of the Analytic Hierarchy Process*