3 Development of the agrobiodiversity management yardstick

Mariette van Amstel, Willem de Neve, Joop de Kraker and Pieter Glasbergen. (2007a) 'Assessment of the potential of eco-labels to promote agrobiodiversity'. *Ambio*, 36(X) (in press).

3.1 Introduction

As a signatory to the Convention on Biological Diversity (CBD), the Dutch government has adopted conservation and sustainable use of biodiversity as policy goals. The CBD defines the in-situ conservation as "the conditions where genetic resources exist within ecosystems and natural habitats, and, in the case of domesticated or cultivated species, in the surroundings where they have developed their distinctive properties" (article 2 CBD). According to the CBD sustainable use means "the use of components of biological diversity in a way and at a rate that does not lead to the long-term decline of biological diversity, thereby maintaining its potential to meet the needs and aspirations of present and future generations" (article 2 CBD). These goals also concern the on-farm biodiversity on private lands (agrobiodiversity), as these covers about 70% of the surface area of the Netherlands (LNV, 2004; Berkhout & Van Bruchem, 2003). Moreover, there is a growing awareness that agriculture and biodiversity are interdependent (Wood and Lenné, 1997; 1999). The CBD lists several strategies for conservation and sustainable use of biodiversity. One such strategy is to encourage the private sector to develop methods for sustainable use of components of biodiversity (article 10 CBD), which may be achieved through industry self-regulation (Gunningham and Grabosky, 1998; Dedeurwaerdere, 2002; Gunningham and Young, 2001).

An eco-label is a type of self-regulation of the agro-food chain that may have potential for conservation and sustainable use of agricultural biodiversity (GTZ, 2001; WBGU, 2001). To assess whether an eco-label is an appropriate governance strategy, a major question to be answered is: in what way and to what extent do current eco-labels stimulate conservation and sustainable use of agrobiodiversity? By answering this question, this paper aims to contribute to the development of knowledge to assess the role of self-regulation in conservation and sustainable use of agrobiodiversity.

Currently, it is not clear to what extent conservation and sustainable use of agrobiodiversity is stimulated by eco-labels. Some eco-labels claim to promote agrobiodiversity (Manhoudt et al., 2002; De Snoo and Van de Ven, 1999), while others seem to avoid this complicated issue diplomatically. Many actors – especially those from the agro-food chain – consider the concept of agrobiodiversity as too abstract, unclear, inaccessible and immeasurable. Although several actors felt that many environmentally friendly regulations already promoted agrobiodiversity in the Netherlands, none could give an overview of all existing regulations and an explanation how these regulations contribute to conservation and sustainable use of agrobiodiversity (Van Amstel et al., 2003). Thus, for an assessment of the potential of eco-labels to stimulate conservation and sustainable use of agrobiodiversity we need an instrument that enables us to link the content of regulation explicitly to agrobiodiversity. In developing this instrument, we faced two major challenges: how to make the abstract concept of agrobiodiversity operational, and how to deal with current knowledge gaps on the relationships between specific agricultural practices and agrobiodiversity?

Our solution to the first problem was to introduce a "ladder of abstraction" (Sartori, 1991). A ladder of abstraction distinguishes several levels of abstraction and makes it possible to "descend" from abstract agrobiodiversity policy goals to the concrete level of management measures on a farm.

The second problem concerns the fragmentary, location specific, often uncertain and partly inconsistent nature of the present published ecological knowledge on agrobiodiversity (Kassas, 2002; Struik and Almekinders, 2000). This makes it impossible to base the instrument purely on scientifically proven and published relationships between management measures and agrobiodiversity. Therefore, we decided to strive for plausibility based on expert judgment. Hence, the resulting instrument, called the "agrobiodiversity management yardstick" (AMY), is underpinned by literature reviews, expert interviews and, most importantly, an expert consensus workshop.

In this paper, we will first describe the development of the yardstick and subsequently its application to five eco-labels. The yardstick is developed for and applied to Dutch eco-labels in arable farming. However, as these eco-labels are typical representatives of international labeling families, the results of our assessment are presumably relevant to other countries as well.

3.2 Development of the yardstick

The relationship between an eco-label and agrobiodiversity can be described as a chain with four elements (box 3.1). Ideally, the potential of eco-labels to promote agrobiodiversity would be assessed with a comprehensive field study targeting the fourth element of the chain. However, even apart from the many methodological problems, a comparative assessment of this kind of five eco-labels would be extremely time-consuming and costly. In our study, we focused on the second element of the chain, the labeling schemes. Thereby we assumed that full compliance with the standards of the labeling schemes is guaranteed by the certifying institutions. These standards may vary in the extent to which they are compulsory, however, and this aspect is included in the study. The eventual impact of the standards on agrobiodiversity was not measured, but assessed indirectly with a yardstick for agrobiodiversity management. This yardstick is basically a checklist of on-farm management measures that contribute to conservation and sustainable use of agrobiodiversity.

The extent to which eco-labels stimulate these policy goals can thus be evaluated, by comparing the standards of the labeling schemes with the measures of the yardstick.

The yardstick was developed in three steps: (a) A study of scientific literature on relationships between agricultural practices and agrobiodiversity. (b) Interviews with four experts discussing

Box 3. 1 Relationship between an eco-label and agrobiodiversity.

Eco-label. An product eco-label sets certain goals concerning reduction of environmental burden and/or stimulation of positive environmental impact of a production process. These goals may include promotion of agrobiodiversity.

↓

Labeling scheme. A labeling scheme is a set of standards an eco-label holder (producer) has to comply with. For each product or product group different labeling schemes are designed.

↓

Standards. The British Standards Institution and the International Organisation for Standardisation (ISO) define a standard as: "a published specification that establishes a common language, and contains a technical specification or precise criteria and is designed to be used consistently, as a rule, a guideline or a definition" (www.standardsglossary.com, consulted November 22, 2006). The standards may encourage conservation and sustainable use of agrobiodiversity.

↓

Compliance. The eco-label holder must comply with the standards of the labeling scheme. There is an inspection of compliance by an (independent) third party. Compliance thus concerns the extent to which the standards regarding conservation and sustainable use of agrobiodiversity are met in actual farming practice.

↓

Impact on agrobiodiversity. The measures (step 3) can have a more or lesser positive impact on biodiversity on farmlands.

first ideas for a yardstick. (c) A workshop with a dozen experts to discuss a prototype, resulting in the final version of the yardstick. For the interviews, experts with different agrobiodiversity research areas were selected: an expert about on-farm biodiversity indicators, about soil biodiversity, about organic agriculture and about agri-environmental schemes. About fifty natural scientists, working in the field of agrobiodiversity, were invited for the expert consensus workshop. The invited scientists were involved as researchers or conference visitors in Biodiversity Stimulation Programme of the Netherlands Organisation for Scientific Research or added to the list using the "snowball-method".

The first step was to make an inventory of agrobiodiversity management measures. In this inventory we aimed to cover the full range of arable farming practices that affect biodiversity. Further criteria in our selection of management measures were exclusiveness (no overlap), suitability for eco-labeling schemes, practicality for the farmer, and relevance for Dutch agro-ecological conditions. The literature study resulted in an inventory of over thirty classes of farming activities with a potential impact on agrobiodiversity. In the second step this inventory was discussed with four scientific experts on agrobiodiversity, specializing in soil biology, organic farming, landscape ecology and farm nature development, respectively. On the basis of both the literature study and the interviews with the experts a prototype of the yardstick was drafted (Van

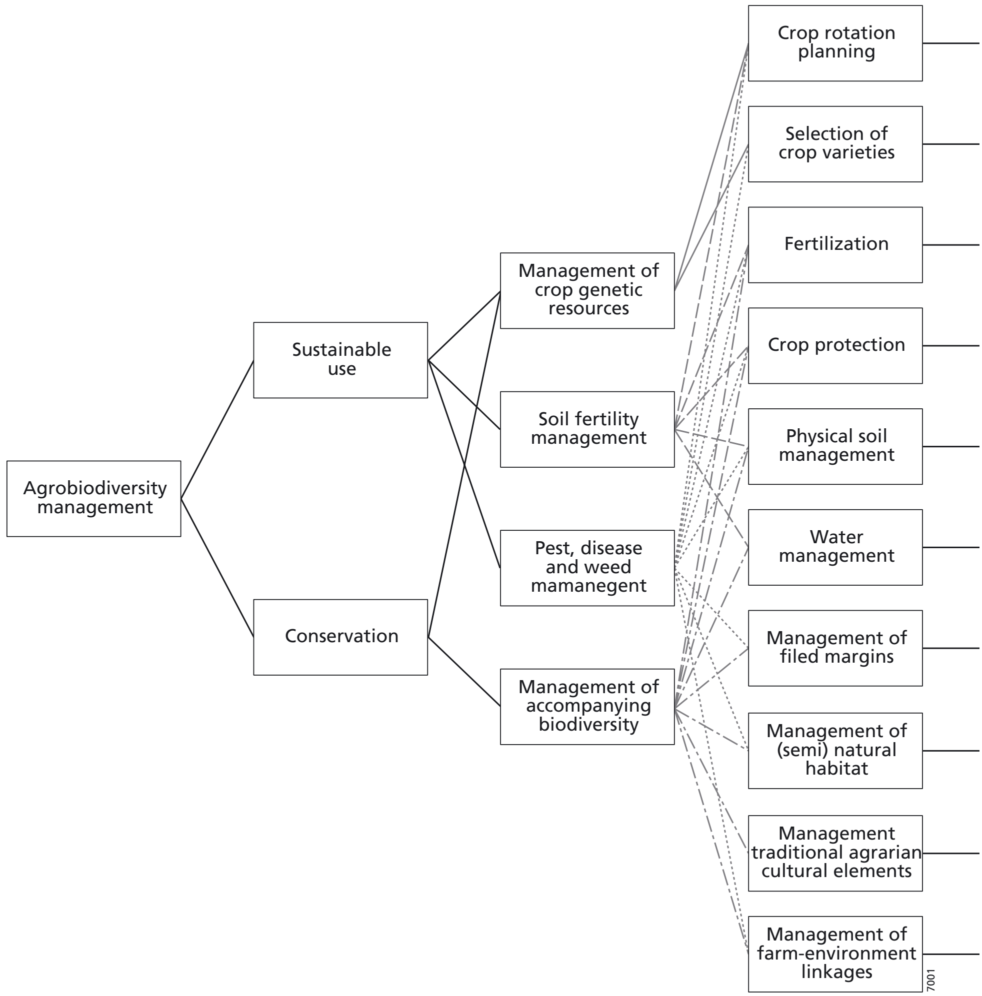

Figure 3.1 The first four levels of abstraction of AMY.

Amstel & De Neve, 2005). This prototype yardstick had five "levels of abstraction", similar to the final version, presented in figure 3.1.

This figure illustrates the first four levels of abstraction of AMY and the relation between the categories of the different levels. The first and highest level of abstraction just contains one category, "agrobiodiversity management". At the second level two major forms of agrobiodiversity management are distinguished, conform the goals of the CBD: "sustainable use of agrobiodiversity with a production function", and "on-farm conservation of agrobiodiversity". The third level emphasizes the different functions of agrobiodiversity management, according to the themes formulated in Dutch national policy on agrobiodiversity (LNV, 2004; VROM, 2005). Agrobiodiversity at stake in the three categories of management of "genetic resources", "soil fertility" and "pests and diseases", is usually referred to as functional biodiversity, i.e. biodiversity with an agronomic production value (Oerlemans et al., 1999). The fourth category, "accompanying biodiversity", does has a production function of public goods and contributes to landscape and recreational values. At the fourth level of the yardstick ten categories of management activities at farm level are distinguished that may affect the agrobiodiversity management functions at the third level of yardstick.

The fifth level of the yardstick contains about 140 on-farm management measures which impact positively on agrobiodiversity. The measures were extracted from the literature, mainly from Boer et al. (2003a, 2003b), Landschapsbeheer Nederland (1998) and Platform Biologica (2002), or suggested by experts. In general these measures make the ecological conditions on farmland less harsh and more diverse, e.g. by reducing the input of biocides and the disturbance or compaction of soil, providing a greater diversity of food sources and breeding sites, and establishing corridors for dispersal (cf. Asteraki et al., 2004; Cortet et al., 2002; Duelli and Obrist, 2003; Hole et al., 2005; Pokarzhevskii and Krivolutskii, 1997; Schippers and Joenje, 2002; Smart et al., 2005; Stoate et al., 2001). A full overview of these measures can be found in De Neve (2006).

The final version of the yardstick was developed in a consensus workshop with a dozen scientific experts on agro-ecology and agrobiodiversity in the Netherlands. This consensus workshop was held in a group decision room (GDR), with a network of computers. This network facilitates the communication process of conventional discussions (Wealtherall and Nunamaker, 1999). The prototype yardstick was discussed level by level in several rounds, each consisting of a plenary introduction to the topic, parallel (electronic) collection of the responses of the experts, plenary discussion of the responses, and parallel assessment of the outcomes by the experts, with the electronic voting system. The four functionally defined management themes at level 3 were generally accepted and only adapted in wording. Discussion and assessment of the categories of farming activities at level 4 resulted in the addition of two new categories (water management and management of farm-environment linkages), and the removal of a few categories which received scores lower than 6 on a 10-point scale. The relationships between the categories at level 4 and the functional themes at level 3 were also assessed by the experts (0 = no relation, 1 = relation). Only relationships with a mean score larger than 0.5 were maintained. At level 5, numerous additional management measures were suggested by the experts. Subsequently, all management measures were assessed with regards to their efficacy. Efficacy refers to the positive impact of a management measure on agrobiodiversity, the extent to which it contributes to conservation and sustainable use of agrobiodiversity. Efficacy was scored individually by the experts on a five-point scale, ranging from 1 (= this measure is not effective at all) to 5 (= this measure is certainly effective). Only measures with a mean score larger than 2.0 were included in the yardstick. In table 3.1 examples of such measures are presented, along with the mean efficacy score of the management measures per farming activity. Category 10, "management of farm-environment linkages" is a cross-cutting management activity, and many concrete measures to tune the farm to its environment are grouped under the preceding activities (e.g. category 6, table 3.1). The measures that are included in category 10 generally pertain to formal-legal aspects of management of farm-environment linkages, such as an obligation for a conservation plan. The complete list with management measures can be found in De Neve (2006) in Dutch. About two-third of management measures were identified in the farming categories "management of field margins", "management of semi-natural habitat", "management of traditional agrarian cultural elements" and "management of farm-environment linkages".

Table 3.1 Mean efficacy scores of experts attributed to management measures per farming activity (AMY level 4) and examples of management measures with efficacy scores above and below average.

Categories of farming activities and examples of management measures	Efficacy score
1. Crop Rotation Planning	3.30
- On adjacent fields, crops are grown from different product groups. E.g. wheat and barley are not grown on adjacent fields, but wheat and sugar beet are.	3.70
- Cropping frequency in the rotation is 1:6 years.	3.00
2. Selection of Crop Varieties	3.44
- Regional varieties are grown.	3.90
- Growing of rare varieties is standard practice.	2.90
3. Fertilization	3.55
- Compost is applied, (partly) replacing fertilizer.	4.00
- Animal manure is applied in place of inorganic fertilizer.	2.90
4. Crop Protection	3.60
- Non-crop vegetation attracting natural enemies is established on the farm.	4.29
- Use of mechanical weed control is maximized.	3.00
5. Physical Soil Management	3.71
- The longevity of permanent grass land is maximized.	4.83
- In the fields, always the same wheel tracks are used.	3.33
6. Water Management	3.00
- Water drainage of farmland is in tune with surrounding nature reserves.	3.83
- The farm does not participate in a project for temporary water storage on farmland.	2.33
7. Management of Field Margins	3.39
- Field margins cover at least 5% of the farm land.	4.67
- Field margins are planted with green manure crops.	2.00
8. Management of (Semi) Natural Habitat	4.03
- Natural waterways, such as a brook, are maintained.	4.83
- Species-specific management is implemented for at least two mammal species.	3.33
9. Management of Traditional Agrarian Cultural Elements	3.52
- A traditional farm orchard (with widely spaced, tall fruit trees) is maintained.	3.80
- When harvesting hay-producing crops, hay is stored in a traditional haystack.	2.60
10. Management of Farm-Environment Linkages	3.68
- The farm is a member of an agri-environmental association with an approved, regionally coordinated management plan.	4.29
- The farm participates in a research project on (functional) biodiversity.	3.14

3.3 Assessment of five eco-labels with the yardstick

Selection of eco-labels

All major institutionalized Dutch third-party product eco-labels for arable farming were included in the study. International, European, national, and regional labels were part of the selection. The national and regional eco-labels have comparable counterparts in other countries, which make them internationally relevant as well. All these eco-labels are, to a more or lesser extent, guided by the European Norms NEN-EN 45011 on product certification systems and the

ISO 65 Guide of the International Organization for Standardization (ISO, 1996). The following eco-labels were selected:

- *EKO*. Since 1992, as a result of the European Regulation on Organic Farming EEC no. 2092/91, The Netherlands have one single organic eco-label, EKO, which is owned by Skal, a private foundation (SKAL, 2005).
- *Demeter*. The Dutch biodynamic (BD) association, founded in 1937, is a member of Demeter International. This BD eco-label includes the standards of EU regulation no. 2092/91. It also sets conditions that go beyond the scope of this EU regulation (Vereniging voor Biologisch-dynamische Landbouw en Voeding, 2004).
- *Erkend Streekproduct, ESP* [Recognized Regional Product]. The foundation Streekeigen Producten Nederland (SPN) [Regional products of the Netherlands] has established several principles and framework standards emphasizing regional aspects and sustainable production (Stichting Streekeigen Producten Nederland, 2000). Regional organizations have elaborated the national standards in response to the regional situation.
- *EurepGAP* is the abbreviation for the Euro-Retailer Produce Working Group (Eurep) combined with the acronym for Good Agricultural Practices (GAP). It is an initiative of several European retailers founded in 1997, that has evolved into a global partnership between agricultural producers and their retail consumers for safe and sustainable agriculture (EurepGAP, 2004a-d).
- *Milieukeur* [Environmental Label]. Since 1995, SMK [formerly Stichting Milieukeur: Foundation environmental label] has owned and developed this eco-label (SMK, 2004 2005a-b). The labeling schemes are based on Life-Cycle Analysis, following the European Eco-label for non-food products. The International Organisation for Standardization has standardized Life-Cycle Analysis within the ISO 14040 series (ISO, 1997).

Assessment of eco-labels

The standards of eco-labels, as defined in box 3.1, were assessed in several ways. The number of standards per labeling scheme contributing to conservation and sustainable use of agrobiodiversity was determined by comparing and contrasting these standards to the management measures in AMY (level 5). By clustering the relevant standards per farming activity (AMY level 4), the width of coverage of the eco-label was determined. Furthermore, the average efficacy of the relevant standards of a labeling scheme per farming activity was determined, and lastly, the compulsory nature of these standards was examined.

The comparison of the labeling schemes with AMY yielded three types of standards relevant for agrobiodiversity:

- Type I: standards matching management measures in AMY (circa 50%).
- Type II: substantive standards that did not match any measures included in AMY level 5, but judged by us to have a positive impact on agrobiodiversity (circa 30%). Some of these standards are formulated in more general terms and match level 4 of AMY. The farmer is given a responsibility and choice how to interpret and adopt more sustainable measures for that particular farming activity. Some standards are related to management measures in AMY, but do not match closely. Others are formulated more indirectly: for example requirements to machinery for crop protection resulting in one of the level 5 management

measures. Finally, there is a small number of standards that were not mentioned by the literature or the experts (circa 6 %).

- Type III: registration standards (circa 20%). The registration standards are procedural and regulate the obligation to keep records of the activities on the farm. These standards are relevant because they create a possibility to check farmers' compliance with the substantive standards. The impact of these criteria is indirect; the direct impact on agrobiodiversity is caused by substantive standards aiming at physical measures.

Further analyses were only conducted with type I and II standards, type III standards were excluded due to their procedural character.

After identifying and counting the number of type I and II standards in each labeling scheme, the standards were clustered per farming activity and their mean efficacy was calculated. For type I standards the efficacy scores were used of the measures in AMY level 5 to which the standards were matched. For type II standards, however, such scores were lacking, as they could not be matched to any of the measures in AMY rated by the experts in the consensus workshop. Therefore, we attributed to Type II standards the mean efficacy score of the farming activity to which they belonged (table 3.1), minus 1 point. The reason for this downward correction was that type II standards had neither emerged from our literature study, nor were they mentioned spontaneously by the experts. Therefore they are expected to be less effective in promoting agrobiodiversity. The exact value of the correction factor (minus 1 point, on the five-point scale for efficacy) is arbitrarily chosen. The mean efficacy per farming activity was calculated for each eco-label as the arithmetic mean of the efficacy scores of all type I and II standards recorded in that category.

In addition to this, the nature of the standards of labeling schemes has been examined. Standards can be compulsory or optional, or they take the form of recommendations. Farmers are obliged to comply with compulsory standards. Typical of compulsory standards are the expressions "must", "is not allowed", "totally forbidden" or "any use of not allowed means leads to decertification". Optional standards manifest themselves in various ways. An example is a threshold criterion. This entails an obligation to comply with a certain number of standards. Examples of expressions that are used by optional standards are "ought to…or….", "minimal 1 measure of the following list" or "requirements must be formulated (content of these requirements are free applicable)". Recommendations demand voluntary compliance with standards. A farmer can choose to comply or not, but it has no consequences. Examples of expressions that are used for recommendations are "attention needs to be paid to", "is recommended" and "need to be aimed for". An overview of all expressions considered characteristic for the three categories can be found in De Neve (2006).

Results of the assessment

All of the examined labeling schemes contain standards that stimulate conservation and sustainable use of agrobiodiversity, but there are large differences between the eco-labels (figure 3.2). The number of type I and type II standards ranged from 6 (ESP) to 42 (Milieukeur).

When type III standards (on registration) are included, EurepGAP has the highest score with 63 standards (not shown in figure 3.2). This large number of registration standards is a consequence of EurepGAP's emphasis on traceability and transparency. The registration standards specify in detail what and how to register. All eco-labels contain standards in at least 6 of the 10 categories

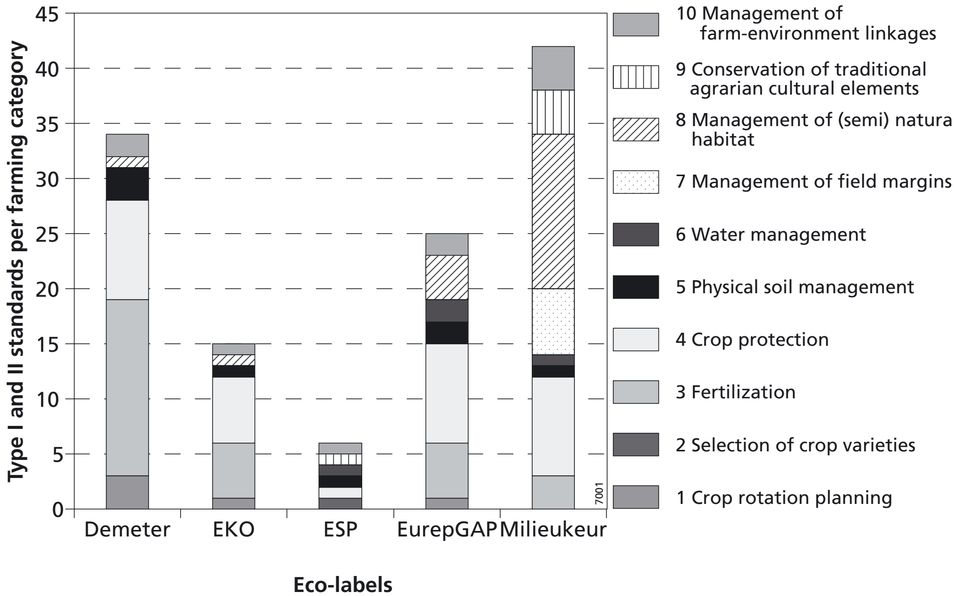

Figure 3.2 Number of standards per farming activity, per eco-label

of farming activities, but none of the eco-labels covers all these categories. Demeter, EKO and Erkend Streekproduct score on 6 categories of farming activities, EurepGAP on 7, while Milieukeur has 8 categories in the labeling scheme. When only type I standards are considered, EKO and Erkend Streekproduct score in 3 categories. The type I standards of EurepGAP are found in 4 categories, the Demeter type I standards in 5 categories and Milieukeur type I standards in 8 of the 10 categories of farming activities. Several eco-labels have only 1 or 2 standards per farming activity. All labeling schemes, except for the one of Erkend streekproduct, contain type I and type II standards in the categories of fertilization and crop protection. This means that all these labels have standards that limit the possible negative impact of crop protection and fertilization on biodiversity. Standards in the categories of soil management and management of farm-environment linkages were found in all labeling schemes. Yet within the four categories mentioned, the eco-labels usually differ in the number of standards and the management measures these standards represent. Standards in the categories of management of field margins and selection of crop varieties were only represented in the labeling schemes of Milieukeur and Erkend streekproduct, respectively.

The mean efficacy of the type I and II standards found in the labeling schemes varies substantially among the eco-labels in most categories of farming of activities (figure 3.3). The differences however between the eco-labels are not consistent. There is no eco-label that has the most effective standards for agrobiodiversity management in all categories of farming activities.

Compared to the means per farming activity (table 3.1), the means of the standards of eco-labels in that particular category are lower most of the time. This is not due to the lower means of type II standards, but for almost all eco-labels the means of type I standards per farming activity

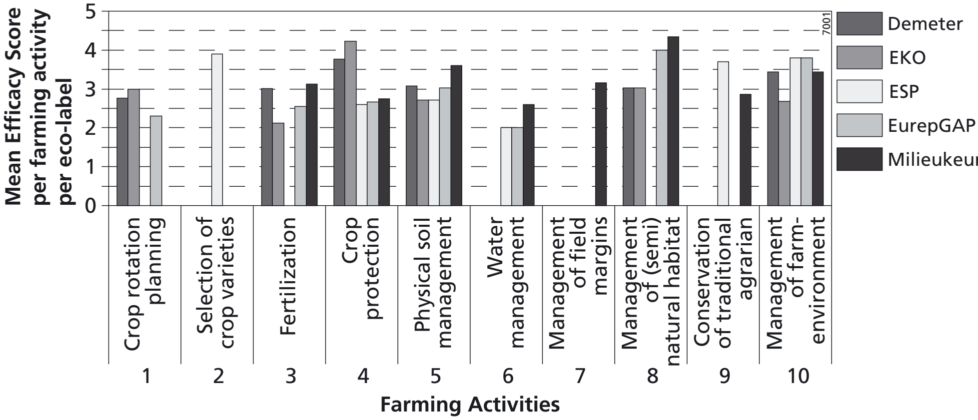

Figure 3.3 Efficacy of standards per farming activity, per eco-label. See table 3.1 for explanation on the numbered categories of farming activities. 1 – this measure is not effective at all. 2 – this measure is not very effective. 3 – this measure is effective. 4 – this measure is very effective. 5 – this measure is certainly effective as a means of agrobiodiversity management.

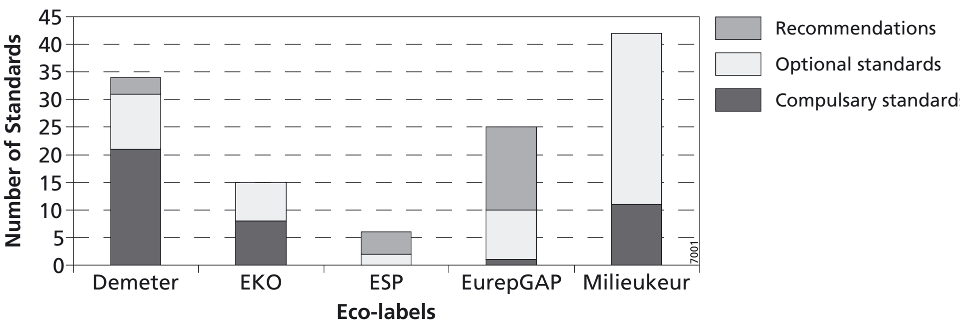

Figure 3.4 Nature of standards, per eco-label

are lower than the means of level 5 management measures for the same farming activity of AMY. This indicates that in general the labeling schemes of the eco-labels contain measures that were judged by the experts as relatively less effective for agrobiodiversity management compared to the other management measures in AMY. Exceptions are Erkend Streekproduct in category 2 and 9, Milieukeur and EurepGAP in category 8 and all eco-labels in category 10.

When the nature of the type I and II standards in the labeling schemes is considered, large differences between the eco-labels become apparent (figure 3.4). In EKO and Milieukeur all relevant standards are either compulsory or optional, whereas in EurepGAP most of the type I and II standards are only recommendations. Demeter is the eco-label with most compulsory standards (20), followed by Milieukeur (11). The biodynamic label Demeter includes all the compulsory and optional standards of the organic EKO-label, but the additional requirements result in a doubling of the number of standards that are agrobiodiversity-friendly.

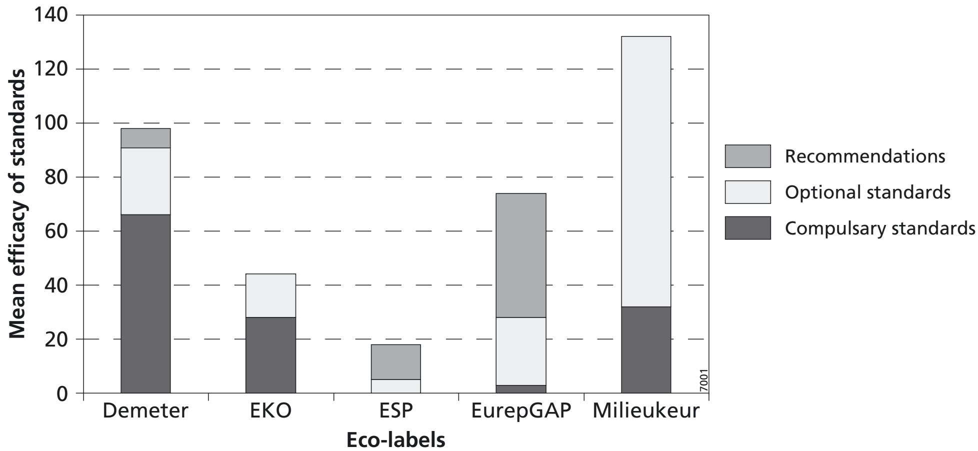

Figure 3.5 Number* mean efficacy of type I and II standards, per eco-labels. (Fig.3.2) multiplied by means efficacy (Fig. 3.3)

An integrated picture of the results is provided by figure 3.5, where the total number of type I and II standards per eco-label is "weighted" for efficacy, by multiplying the number of standards per farming activity (figure 3.2) with the corresponding calculated mean efficacy (figure 3.3). The maximum attainable value on the Y-axis would be 595 (175 * 3.4), which is highest recorded number of standards for an eco-label multiplied with the mean score for efficacy of all 175 measures. Given this maximum, it can be concluded that all five labels have a limited efficacy and can improve in absolute terms. Figure 3.5 shows that Demeter has the highest combined score of compulsory measures and efficacy for agrobiodiversity management, but that Milieukeur has the greatest potential for agrobiodiversity management due to the many optional standards with a high efficacy score.

3.4 Discussion and conclusions

Approach

The yardstick for agrobiodiversity management is designed as a ladder of abstraction with five different levels. It proved possible to descend on the ladder to the level of concreteness needed to assess the labeling schemes of eco-labels and still have a plausible relation with the top of the ladder. The ladder of abstraction made it possible to compare how and to what extent different eco-labels stimulate conservation and sustainable use in their labeling schemes for arable farming.

To construct the yardstick, ecological expert knowledge about agrobiodiversity was mobilized through a consensus workshop. This proved necessary, especially for the construction of the fifth, most concrete level of the yardstick, because from the scientific literature a consensus view cannot be derived. The knowledge about agrobiodiversity management is fragmentary and sometimes inconsistent, and the focus in the scientific literature is on novel or unique findings and disagreements, rather than on consensus and agreement. Also during the consensus workshop, the discussions on the yardstick prior to assessment by voting showed a rather strong divergence, with a focus on disagreements. Interestingly, in the (parallel) assessment it turned out that the experts disagreed only on circa 10% of the measures that were discussed, while they agreed on circa 90%. The expert consensus workshop was thus an essential tool to provide us with a clear image of what Dutch scientific experts agree upon with regards to the relationships between farming practices and agrobiodiversity.

The order in which the yardstick has been developed could have been different. We first developed the yardstick, and then analyzed the eco-labels. As a consequence we had to distinguish the type II standards in eco-labels: the standards that possibly affect agrobiodiversity, but are not included in AMY. The effect of these standards on agrobiodiversity is less certain, because the experts did not judge their efficacy. It would be tempting to conclude that the order of the study should be reversed: first analyze the eco-labels and then have the experts judge them. That would however cause two problems. First, it is methodologically more correct to develop a yardstick and then measure with it, since the reverse approach would (unconsciously) influence the results. Another, even more urgent reason is that the yardstick would develop a bias to these five eco-labels, which is not desirable since we intend to apply AMY also to other types of regulation. Therefore we chose for a flexible and open mechanism on level 5 of AMY, creating the opportunity to add the other – type II – measures, even though it is more laborious.

Another issue is the objectivity of the assessment method. The decision as to whether a standard of an eco-label matches a management measure of AMY can sometimes be rather subjective. There is a tension between a broad or a narrow formulation of a management measure in AMY. The disadvantage of a broad formulation is that different interpretations are possible, whereas a narrow and accurate formulation may make it difficult to match the standards of eco-labels with the management measures of the yardstick. The formulation of the standards differs among the eco-labels and differs from the formulation in the yardstick. Some standards are formulated in a way that they match several management measures of AMY. A choice has to be made to which management measure the standard is matched. As mentioned above, we have chosen a rigid interpretation and have designed the category of type II standards for the standards that did not meet the management measures fully. The choice for an open level 5 of AMY is connected with this more rigid interpretation of the (type I) standards.

Despite these methodological issues, the yardstick for agrobiodiversity management proved to be useful for the purpose for which it was developed. In absolute as well as comparative terms we could assess with AMY how and to what extent labeling schemes of five current eco-labels stimulate conservation and sustainable use of agrobiodiversity. AMY may be utilized for assessment of eco-labels of other agricultural sectors and other regions as well, although not as a ready-made yardstick. Useful elements are the ladder of abstraction and the consensus workshop to substantiate the yardstick with expert knowledge. Furthermore, the large amount of management measures at level 5 will provide a useful starting point, but of course adaptation is required according to the specific nature of the agricultural sector, such as horticulture and cattle breeding, or to the agronomic, geographical and climatological conditions of a particular region or country. In addition to eco-labels, AMY may be applied for the assessment of other types of (self-) regulation for conservation and sustainable use of agrobiodiversity. In such an assessment, AMY could be used for evaluation of governance mechanisms like private regulation and policy instruments.

Results

The extent to which the five eco-labels stimulate agrobiodiversity management is in general limited and differs substantially among the labels. EKO and Demeter distinguish themselves by addressing the core activities of the farmer. They focus in particular on fertilization and crop protection – two farming activities that are often considered to have a negative impact on agrobiodiversity. Milieukeur can be characterized as the eco-label that emphasizes conservation of biodiversity on the non-productive areas of the farm (field margins, (semi) natural habitat and traditional farm elements), whereas Erkend streekproduct emphasizes the regional aspects (regional crops and varieties) of agrobiodiversity. For EurepGAP both the fertilization and crop protection are relevant categories, but also the conservation of (semi)natural habitat. None of the eco-labels included in our study refers explicitly to agrobiodiversity in their goals. In two cases, Erkend streekproduct and EurepGAP, the word (agro) biodiversity was mentioned in the labeling schemes, but despite of this these two rank lowest among the five labels in terms of number of compulsory or optional standards relevant to agrobiodiversity.

Compared to the four other eco-labels, Erkend streekproduct stimulates least the conservation and sustainable use of agrobiodiversity. The number of relevant standards in its labeling scheme is considerably lower than in the schemes of the other eco-labels, and although the standards are unique, none of these are compulsory. This can be explained by the function of Erkend streekproduct, which is intended as a framework for smaller regional labels, which set their own specific standards.

Of the remaining four eco-labels EurepGAP is the weakest, because most standards are optional or recommendations. Milieukeur and Demeter appear to stimulate agrobiodiversity management most. The labeling scheme of Milieukeur contained the highest number of standards relevant to agrobiodiversity (42), with the best coverage of farming activities (8 categories). The broad coverage of farming activities by Milieukeur can be explained by its basis in Life Cycle Analysis, which aims to include the environmental aspects of the complete production chain. The relative weakness of Milieukeur is that about 75% of these measures are optional. Demeter is the eco-label with the highest number of compulsory biodiversity-friendly measures (20).

The ranking of the eco-labels is a relative measure. In absolute terms, none of the five eco-labels had a high score. The scores varied from 6 to 42 standards per eco-label, out of the circa 175 on-farm measures in AMY (type II standards included). In all cases the score included optional standards and recommendations, which makes compliance of farmers with all of these standards less likely. An interesting outcome of this study is that all ten categories of farming activities are covered when the results of the five eco-labels are combined. This means that eco-labels could learn from each other and improve by exchanging knowledge and experiences with standards relevant to agrobiodiversity.

Conclusions

Are eco-labels an appropriate instrument for conservation and sustainable use of agrobiodiversity? AMY shows that it is possible to translate the abstract, comprehensive concept agrobiodiversity management into concrete on-farm measures. These management measures, sometimes formulated at the higher level of abstraction of the farming activities, are found in eco-labels. This makes eco-labels in theory an appropriate instrument for conservation and sustainable use of agrobiodiversity. In practice however, management measures for agrobiodiversity are as yet not very well represented in the labeling schemes of eco-labels. Even the highest-ranking eco-labels do not have a substantial number of compulsory standards in all categories of farming activities: Demeter has a relatively narrow focus on fertilization and crop protection, whereas the standards of Milieukeur are optional in many cases. Compliance to these eco-labels does therefore not guarantee the agrobiodiversity-friendliness of all relevant farming activities. This is a major drawback, as all these aspects of farm management are interacting factors in sustaining agrobiodiversity, and their joint effect is likely to be as strong as the weakest link.

Can we expect an "improvement" of eco-labels as an instrument for conservation and sustainable use of agrobiodiversity? We do not expect substantial improvements in the near future because agrobiodiversity is not a prominent theme in societal debate and government policy (Van Amstel et al., 2005). Especially for those eco-labels who now have improved compliance with the standards as a major priority, main changes of the labeling schemes cannot be expected soon. In case of changes, however, AMY could function as a source of inspiration. Another option is the establishment of a new eco-label with a special focus on conservation and sustainable use of agrobiodiversity, but given the multitude of eco-labels already on the market, we do not expect a broad societal basis for such a label.

In addition to the Netherlands, the findings of our study are also relevant to other countries. The standards of Demeter and EurepGAP are applied worldwide, whereas EKO is based on EU regulation for organic farming. Milieukeur and Erkend Streekproduct are part of international labeling families based on Life-Cycle Analysis (ISO 14040, 1997) and protection of regional products. The outcomes of our assessment for the Netherlands are thus probably a good indication of the potential of eco-labels for self-regulation of conservation and sustainable use of agrobiodiversity elsewhere.