MÉSZÁROS Dóra*[1], HUFNAGEL Levente*[2], BALÁZS Katalin*[3], BÍRÓ Zsolt*[4], JANCSOVSZKA Paulina*[5], PODMANICZKY László* and SIPOS Balázs*

Farm-level environmental performance assessment in Hungary using the Green-point system

Faced with society's increasing expectations, the European Union's Common Agricultural Policy uses environmental management as an increasingly critical criterion in the allocation of farm subsidies, with a shift in focus from production and area-based subsidies to payments for supplying public goods. There is an increasing demand to assess the ecological and environmental performance of farms as public money spent on provision of environmental services requires justification. The objective of this research is to strengthen the basis of the concept of farm-level environmental performance assessment. Firstly we give an overview of indicator-based sustainability assessment tools. Even though there are several different tools developed globally, and the themes and indicators for the assessment of environmental performance are very similar, there are significant differences in terms of data survey among them. Secondly we describe the development and field testing of the 'Green-point system' developed in Hungary. This system is able to measure the environmental performance of farms and their value/capability of providing public goods and sustaining ecosystem services through a framework of farm enterprise calculations and assessments. The Green-point system fits well into the stream of yet scarce approaches and efforts, which in several European countries aim to introduce and strengthen the so-called result-based agri-environmental schemes alongside the currently rather dominant management-based approaches.

Keywords: agroecosystems, diagnosis, environmental impact assessment, evaluation methods, indicators, sustainability

* Szent István Egyetem, 1 Páter Károly út, H-2100 Gödöllö, Hungary. Corresponding author: sipos.balazs@yahoo.com; http://orcid.org/0000-0002-7931-9772

Introduction

Agricultural market competitiveness is dependent on its production efficiency and the quality of its products relative to environmental capacity and the status and quality of natural resources. Since the 1960s, intensification of agricultural production has caused increasing environmental pollution, driving much research on the environmental impacts of agriculture (e.g. Wauchope, 1978; Ryden *et al*., 1984). More recently, the loss of biodiversity, soil erosion, deterioration of water quality and the decrease in soil organic matter have received increasing attention (e.g. Pimentel and Kounang, 1998; Kätterer and Andrén, 1999). To protect and enhance the European Union's (EU) rural heritage, the Common Agricultural Policy (CAP) aims to head off the risks of environmental degradation and enhance the sustainability of agro-ecosystems by promoting agricultural practices that preserve the environment by means of a relatively complicated system of regulations and subsidies (Tangermann, 2011).

Conducting farm-level sustainability assessments has several benefits. They provide measurable results and assessment for farmers and also offer the possibility to benchmark farmers against each other once regional anonymous databases are created. They contribute to increasing farmers' awareness of possible environmental improvements and support farm management decisions. Assessment tools can provide baseline information for policy support systems and to result-based agri-environmental schemes. However, there is still relatively little experience with authentic evaluation of environmental success in farming. Development and implementation of suitable assessment tools raise several questions such as what indicators should be used to express agri-environment relationships on a farm in a way that facilitates improved management; how can environmental improvement be documented using appropriate indicators; what indicators have been developed already and how useful are they for farmers and advisors; and how should the set of indicators be defined such that sustainability assessment is the least complex to complete but still provides useful evaluation.

The objective of our work is to strengthen farm-level assessment of environmental performance. Firstly, we give an overview of the importance of indicators as tools for assessing sustainability with a focus on the criteria that make an indicator appropriate for farm-level environmental evaluation. Secondly, we describe the development of the 'Green-point system' indicator set for Hungarian agriculture. Thirdly, we present the results of the farm and field-level testing of these indicators.

Indicators as the basis of assessments

The term 'indicator' has been defined as 'a variable which supplies information on other variables which are difficult to access and which can be used as a benchmark to take a decision' (Gras *et al*., 1989). Indicators should have three dimensions: systemic, temporal/spatial and ethical. *Systemic* means that they are required to assess the economic, environmental and social aspects of agriculture. *Temporal and spatial* indicate the purpose to assess the effects that are likely to occur over time and in space, and *ethical* refers to the sustainability which is founded on a system of values such as the need to conserve the natural and human heritage, or at least to use it as sparingly as possible (Zahm *et al*., 2006).

During the *Results-based agri-environment schemes* conference in 2014 (IEEP, 2014), participants concluded that

[1] http://orcid.org/0000-0003-2882-6683
[2] http://orcid.org/0000-0003-0438-2432
[3] http://orcid.org/0000-0001-7945-1880
[4] http://orcid.org/0000-0002-6046-5833
[5] http://orcid.org/0000-0002-1729-678X

indicators used for this purpose should be (a) clearly linked to objectives (e.g. ecosystems, specific habitats or particular species); (b) reliable (based on adequate evidence) measures of the overall desired outcome, which must be appropriate to context and location; (c) set within a simple framework with common payment triggers – perhaps with two or three hierarchical levels; (d) relatively easy to identify and survey (hence cost-effective); (e) linked to wider goals and user needs (e.g. RDP Monitoring and Evaluation Framework indicators); (f) cheat proof; and (g) acceptable to farmers.

Thus, the development of indicators that can measure the sustainability of an agricultural production system is a complex task. The main question is how to translate the concept of sustainability into operational terms at an individual farm level. During the past 25 years, several different approaches have been developed for assessing aspects of farm-level agricultural sustainability. Pacini *et al*. (2004) used a model to compare the economic, agronomic, technical and environmental results for the MacSharry and the Agenda 2000 reforms. Van Passel *et al*. (2007) constructed an economic model to analyse the impact of managerial and structural farm characteristics on farm sustainability. Others (Bechini and Castoldi, 2009; Thomassen *et al.*, 2009) applied economic and environmental indicators to evaluate farm sustainability. Singh *et al*. (2009) and Binder *et al*. (2010) compared and analysed different sustainability methods by means of literature review. Fumagalli *et al*. (2011) calculated agro-environmental and economic indicators to evaluate farm sustainability and also compared current and alternative management scenarios. Carmona-Torres *et al*. (2014) developed a multifunctional farm-level performance assessment, comparing the current and alternative farming techniques.

However only a few operational applications have been described in the literature (i.e. the Ökopunkte scheme in Degenfelder *et al*., 2005; MOTIFS in Meul *et al*., 2008; DIALECTE in Pointereau *et al*., 2012; MESMIS in Ripoll-Bosch *et al*., 2012; and SMART in Jawtusch *et al*., 2013 and in Schader *et al*., 2014). The development of such tools is considered by many authors (e.g. Hansen, 1996; van der Werf and Petit, 2002; IEEP, 2014) as a method to support the implementation of sustainable agriculture. Each uses a set of indicators to express the degree of environmental impact of a farm based on the use of external inputs in relation to the production and/or the use of specific management practices.

The results of indicator-based assessments should be applicable at several levels: for research purposes, for policy makers and as a source of information for the general public. These different groups have different needs. The research community focuses on the methodologies being internally consistent and the data comparable. Policy makers prefer indicators of sustainable development which are clear, unambiguous and helpful to strategic and applied policy making (Hanley *et al*., 1999). As Meadows (1990) points out, ordinary people need to be informed if their environment and quality of life are deteriorating, about whether this trend is expected to continue, and how such a situation be reversed.

The most important stakeholder group of these sustainability assessments are farmers as they use directly the results of evaluations. As evaluations provide measurable results, farmers can do a year-to-year comparison of their own farming practices from the sustainability point of view or they can benchmark their activities against other farmers. These analyses can lead to better management decisions and can extend farmers' knowledge.

Methodology

Our research focuses on the development of the Green-point system for evaluating the environmental performance of Hungarian farms. This was carried out within the framework of the *Terradegra* project coordinated by the Agricultural Research Centre of the Hungarian Academy of Sciences. The main objective of the *Terradegra* project was to provide IT background and database for research services on the environmental load of agriculture and on the environmental condition of soil for the Soil Degradation Subsystem (SDS) of the National Environmental Information System. For the purposes of the research, farms representing all types of agricultural activities from across Hungary were selected. Data were gathered to monitor the effect of agriculture on the environment by defining the most important indicators of soil degradation. During the development and testing of the indicator system these data were used to calculate the Green-points of farms and to analyse statistically our sample.

Development of the Green-point system indicator set

As there has been no previous experience of evaluating the environmental performance of Hungarian farms, the Green-point system is derived mainly from the French DIALECTE system. However, owing to the fact that French and Hungarian farming practices are different, instead of a simple translation of the French tool, some modifications were required. Adaptation of the DIALECTE system had already started before the *Terradegra* project. The DIALECTE indicator set was tested on different pilot areas on a small number of farms. Based on the results, modifications were suggested by the experts participating in the projects. These modifications were done on three levels: (a) some indicators were modified: by keeping the focus area of the given indicator, the concept of measurement was changed; (b) maximum available points were modified in certain cases; and (c) new indicators were introduced to guarantee more precise measurement of farm sustainability in Hungary.

In the Green-point system each indicator is assigned a maximum score. The results of a surveyed farm are calculated through an algorithm based on the management practice of that particular farm. There are two levels of scoring: the field (or plot) and the farm level. While some indicators are broken down to field level, some others are defined only at the farm level. For example, average field size, and the diversity of crop production are only relevant at the farm level. The farm-level scores are based on field-level scores. Some farm-level indicators were calculated by weighting the field-level results with the size of relevant plots while others (for example indicator 16) were calculated by a different algorithm as described below. By running the collected data through specific algorithms to obtain the score of individual indicators, the total score of the farm in the Green-point system could be derived.

Thus the Green-point system is a quantitative evaluation method applicable for each land-use and farming type. Indicators are used to describe the intensity and the environmental effect of the farm. The system enables the user to compare the performance of farms, or within the same farm, the performance of different economic years. This approach motivates the farmer by focusing on the environmental achievements of the farm. The system is to enhance diversity (at farm and at species level), a minimised use of chemicals and artificial inputs, and the application of management methods that are similar to the traditional and extensive ones.

Field testing of the indicator set

Representative sampling of the farms was done based on the database of the Hungarian Central Statistical Office (KSH) General Agricultural Census (GAC), 2010. In the GAC the KSH recorded several parameters per farm and also categorised these 26,557 farms by applying a representative multiplier (indicating the size and the production volume of the farm). This methodology was worked out by KSH and the data of the categorised farms were passed to the *Terradegra* project team. Following this the project team identified the farm types which were characteristic on the national and (NUTS 3 level) county levels.

As the GAC contains several farming-related parameters, the three most important parameters indicating the farming intensity and the environmental impact of a farm were defined: (a) amount of fertiliser applied (kg ha^{-1}); (b) proportion of area where pesticides were used (%); and (c) amount of organic manure applied (t ha^{-1}). For each of the three parameters the 26,557 farms in the GAC were divided into five sub-categories (Table 1). Based on these parameters a three letter code was attached to each farm, where each code/letter (A-E, F-J and K-O, respectively) referred to one parameter. Using the representative multiplier, it showed us the characteristic farm types at national and county levels. The result of this classification was the identification of the most typical farm types and their proportion in each county, which ensured the representativeness of farm selection.

Data were collected by county experts with different professional backgrounds (soil experts, plant protection experts etc.) using a specially-developed questionnaire. Participants answered simple multiple-choice questions and entered on special data sheets the numerical values of certain parameters that clearly define the intensity of management. These answers defined (directly or indirectly) the results of different indicators. The scale of data collection was the largest homogenous area unit, i.e. a plot in the cases of arable and grasslands. There were two (spring and autumn 2011) surveys on two levels (farm and field). Complete data sets were collected on more than 2600 fields of 260 farms and a database was compiled for all examined indicators for the previous three economic years (2008/09, 2009/10 and 2010/11).

Methods of statistical analysis

Firstly, the contribution of each indicator to the total field- and farm-level scores was calculated. The percentages of farms with zero points and highest points for each indicator were used to illustrate the variability in the environmental performance of the farm sample. In addition, the relative contribution of six indicator groups to the variance in the data, and the impacts of four newly-developed indicators on the total scores, were assessed.

Secondly, owing to concerns about the appropriateness of the weightings given to each indicator in the total score, the scales of the 17 indicators were standardised (the achieved points were divided by the maximum available points). The standardised indicators were evaluated from two aspects. On the one hand, a correlation matrix was applied, and this was followed by a cluster analysis (paired group method). On the other hand, the environmental sustainability of the sample farms was assessed from the scores of the indicator groups. A one-way ANOVA test (with Tukey pairwise post hoc test) was carried out to identify significant differences.

Results

Indicators and their calculation methods

The Green-point system indicator set is composed of 17 indicators with a maximum total score of 90 (Table 2). The set is composed of six groups (A-E): four indicators of nutrient management, two of soil protection, three of natural landscape elements, one of plant protection and water management, three of energy consumption and four of diversity of crop production. Of these 17 indicators, seven indicators remained unchanged from DIALECTE, in six cases the weightings or calculation methods were changed, while four indicators (*length of field boundary with hedge*, *area affected by water management*, *irrigation*, and *external services*) were completely new. Thirteen are field-level indicators, all of which have farm-level versions, and four are farm-level only indicators. A further indicator (*crops cultivated*) is only used as background data for calculating different farm-level indicators.

Table 1: Parameters used to categorise the 26,557 farms in the General Agricultural Census 2010 database.

Amount of fertiliser applied (kg ha^{-1})		Proportion of area where pesticides were used (%)		Amount of organic manure applied (t ha^{-1})	
Value range	No. farms	Value range	No. farms	Value range	No. farms
0	11,854	0	7,138	0	19,401
0< ≥200	4,042	0< ≥50	3,240	0< ≥10	2,850
200< ≥500	7,797	50< ≥100	10,955	10< ≥30	2,579
500< ≥700	1,271	100< ≥200	3,655	30< ≥50	1,158
>700	1,593	>200	1,569	>50	569

Source: own composition

Table 2: The parameters of the Green-point system indicators structured by indicator group.

No.	Indicator name	Unit of measurement	Max points	Level of usage*	Modification
A. Nutrient management					
1	Nitrogen balance	Active substance, kg ha^{-1}	10	Fi+Fa	As DIALECTE
2	Phosphorus balance	Active substance, kg ha^{-1}	10	Fi+Fa	Weighting changed
3	Potassium balance	Active substance, kg ha^{-1}	10	Fi+Fa	Weighting changed
4	Use of organic manure	kg ha^{-1}	5	Fi+Fa	As DIALECTE
B. Soil protection					
5	Winter land coverage	Percentage of total area	4	Fi+Fa	As DIALECTE
6	Non-ploughed areas	Percentage of total area	4	Fi+Fa	As DIALECTE
C. Natural landscape elements					
7	Length of field boundary with hedge	Percentage of total length	4	Fi+Fa	New indicator
17	Average plot size	ha	5	Fa	As DIALECTE
8	Area affected by water management	Yes/no	2	Fi+Fa	New indicator
9	Territory of land elements	Percentage of total area	5	Fi+Fa	Calculation method changed
D. Plant protection					
10	Frequency of pesticide use	Area of application (ha)	6	Fi+Fa	As DIALECTE
E. Energy consumption					
11	Irrigation	Yes/no	2	Fi+Fa	New indicator
12	Fuel consumption	l ha^{-1}	3	Fi+Fa	Calculation method changed
13	External services	Yes/no	1	Fi+Fa	New indicator
F. Diversity of crop production					
14	Crops cultivated	ha	-	(a)	-
15	Proportion of legumes in crop structure	Percentage of total area	0	Fa	Weighting changed
16	Proportion of cereals and maize	Percentage of total area	5	Fa	Weighting changed
18	Diversity of crop structure	Percentage of crop area	10	Fa	As DIALECTE

* Fi = farm; Fa = farm; (a) used indirectly for farm-level calculations
Source: own composition

Field-level indicators

Nitrogen, phosphorus and potassium balances (indicators 1-3) are calculated by subtracting the amount of active substance removed from the amount of active substance applied. One part of the information required for the calculation of the applied active substance is directly available from the famers. The other part requires a calculation based on the type and the amount of fertiliser applied on the field. The per-hectare amount of active substance removed is obtained by multiplying the crop yield with a crop specific N/P/K coefficient. Surplus nitrogen is maximised at 50 kg ha^{-1}, no points are given above this. The maximum score is 10 if the surplus is zero, otherwise the logic of the calculation is that the lower the surplus values, the higher the scores. The algorithm is similar for P and K, but the limit is 30 kg ha^{-1}.

Use of organic manure (indicator 4) compares the amount of active substance applied with manure to the total amount of active substance. To obtain the indicator score, manure quantity and type are needed. The algorithm takes these data and an N/P/K coefficient from a background table to calculate the amount of active substance applied with manure per unit of area. The figures obtained are then compared to the total amount of N/P/K applied (including artificial fertilisers) on the field. To get the final score, the average proportion of active substances (N/P/K) applied with manure is multiplied by the maximum potential score (5 points). When only manure was applied as a fertiliser, the maximum score is automatically given.

Winter land coverage (indicator 5) is scored according to crop type (winter wheat, multi-year lucerne, temporary grassland etc.). Four points are awarded when the whole area has winter coverage and the minimum score is 0 for not having any cover. As regards to *non-ploughed areas* (indicator 6), only direct sowing, set aside and other areas that are not ploughed can be rewarded with 4 points. For *length of field boundary with hedge* (indicator 7), the percentage of total length with hedge is multiplied by the highest possible score (4 points) and then with a multiplier of 1.5; however the final score cannot be higher than 4 points.

Area affected by water management (indicator 8) is a simple yes/no question, with a score of 2 points if there are areas under water management and 0 if there are not. In terms of territory *of land elements* (indicator 9), five points are awarded when less than 10 per cent of the field is covered with landscape features (e.g. small lakes, forest belts etc.), and the maximum score (10 points) is given if the proportion is higher than 10 per cent. For *frequency of pesticide use* (indicator 10) the number of pesticide applications is multiplied by the size of the field[6] and then with a multiplier, which is 1 for herbicides, 2 for fungicides and 3 for pesticides. The interim score is obtained by dividing the sum of sub-results by the size of the field. When the interim score is above 13, the final score is zero. When it is below 13, then the maximum score (6 points) is divided by 15 and multiplied by the interim score to get the final score.

Irrigation (indicator 11) is another simple yes/no question. The indicator scores 2 points if there is irrigation and 0 if there is not. The amount of water per area unit is not considered in this calculation. For *fuel consumption* (indicator 12), the Green-point system divides the fuel consumption on a given field with the total area of the field to obtain the specific (per hectare) consumption. If the result is at least 150 l ha^{-1}, the score is 0; between 100 to 150, the score is 1;

[6] If a pesticide was applied three times on the same field, we calculate with 300 per cent of the territory of that field. Above this we apply the multiplier of 3.

between 50 to 100, 2 points are given; and, for results below 50 litres, 3 points are granted. *External services* (indicator 13) is also a simple yes/no question. The indicator scores 1 point if the farm relies on any external service and 0 if it does not.

For *crops cultivated* (indicator 14), only interim points are given, since these data are only needed for farm-level calculations. For grain crops, maize and legumes the interim score equals the size of the field. When there is an annual or multi-year crop on the field, the interim score is obtained by multiplying the field size with 0.1. When there is undersowing on the area, the multiplier is 0.2; whereas for temporary grassland the multiplier is 0.3.

Farm-level indicators

Although most of the indicators are calculated directly from field-level indicators 1-13 (but weighted by the size of relevant area), an additional four indicators are only applied at the farm level.

Proportion of legumes in crop structure (indicator 15) is calculated from the interim scores of the field-level *crops cultivated* data. Relevant field-level interim results are summed and then divided by the total area of the farm. To get the final score this fraction is multiplied by 5 points. The algorithm is the same for the *proportion of cereals and maize* (indicator 16). Finally, the *average plot size* (indicator 17) is calculated. Where this is above 10 hectares, the score is 0; values between 5 to 10 hectares are reduced by 5 and the result is extracted from the maximum available score (5 points); for values under 5 hectares, a maximum score of 5 points is given.

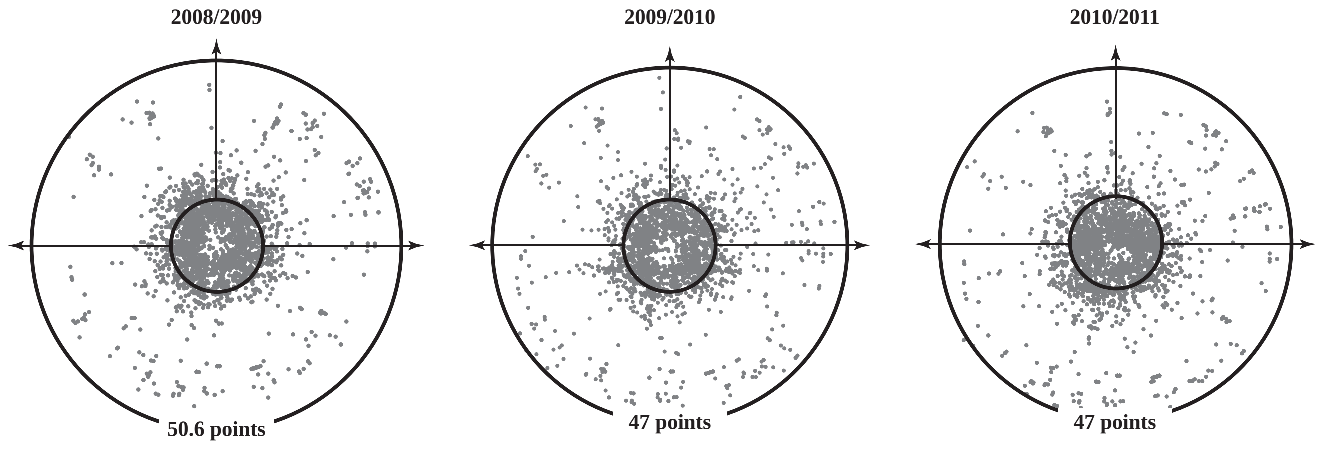
For *diversity of crop structure* (indicator 18), we sum the number of different types of crops cultivated (including grasslands and pastures), and this figure is divided by the number of plots. This is multiplied by the maximum score (10 points). In the second part of the calculation, the percentage of grass coverage is calculated for the total farm size. Multiplying the maximum 10 points with this percentage, the second interim score is obtained. The final score, which should not exceed 10 points, is the sum of these two interim scores.

Field testing of the indicator set

Field results

The field-level scores of the three economic years are shown in Figure 1. The inner circle represents the annual average value, whereas the outer circle shows the maximum value of that year. The maximum possible score at the field level is 60 points, and the field with the best performance reached 50.6, i.e. 84.6 per cent of the maximum. While this particular figure is relatively high, the average values are low, about 23 per cent of the total score. There was no significant difference between the performances of the three years. Probably this is due to the fact that management methods and other relevant parameters are likely to have remained unchanged from year to year.

We examined the influence (weight) and proportion of each indicator within the total score of a particular field and calculated the average importance of each indicator. Four indicators (out of 14) accounted for 76 per cent of the total performance of fields. The indicator with the biggest effect is *frequency of pesticide use* (indicator 10), which accounted for 39 per cent of the total score, followed by *fuel consumption* (indicator 12), *non-ploughed areas* (indicator 6) and *length of field boundary with hedge* (indicator 7) which accounted for 16, 12 and 10 per cent respectively. The first three of these indicators were drawn from DIALECTE (although in the case of indicator 12 the calculation method was changed), while the fourth is a completely new indicator developed by us. The remaining ten indicators accounted for less than a quarter of the total performance of fields.



2008/2009	2009/2010	2010/2011
Number of fields: 2665	Number of fields: 2663	Number of fields: 2665
Maximum: 50.6	Maximum: 47.0	Maximum: 47.0
Minimum: 1.0	Minimum: 0.6	Minimum: 1.0
Average: 13.8	Average: 14.5	Average: 14.0
Deviation: 9.7	Deviation: 10.0	Deviation: 9.7
Relative deviation: 70.6%	Relative deviation: 69.0%	Relative deviation: 69.5%

Figure 1: Green-point field-level scores for three economic years.
Source: own data

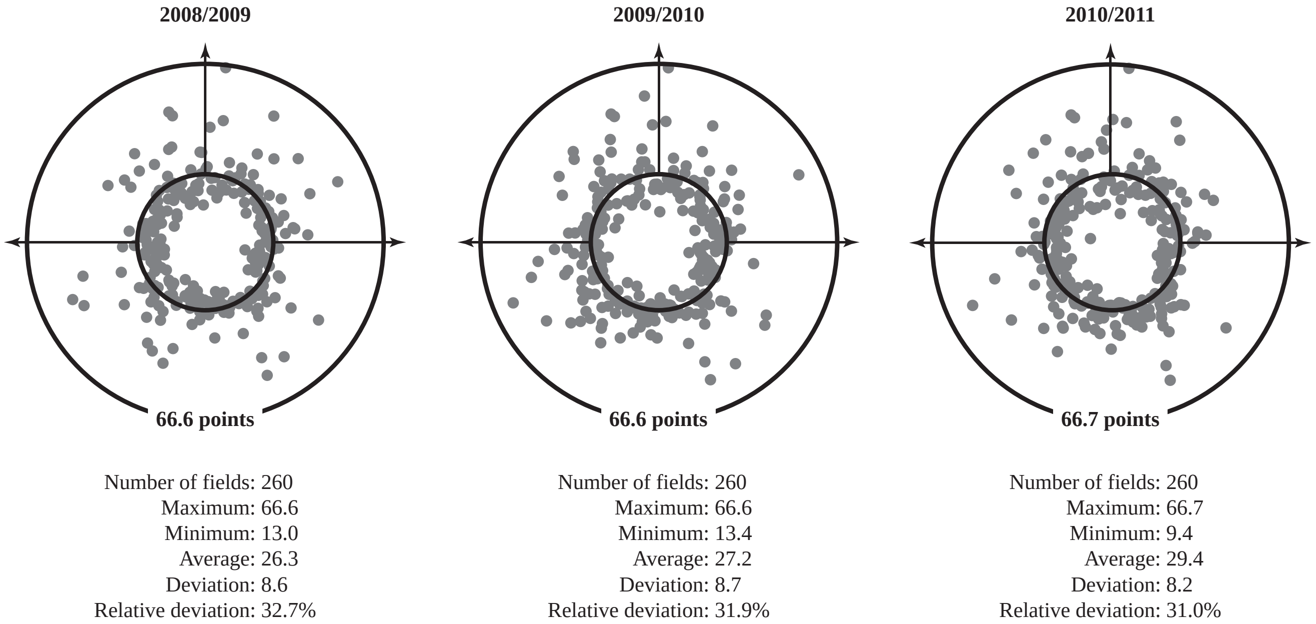


Figure 2: Green-point values of three economic years at farm level.
Source: own data

Farm results

At farm level, the pattern of distribution and the difference between economic years at farm level are similar to those shown on field level (Figure 2). The maximum possible score at farm level is 90 points. The average score was 33 per cent, and even the maximum scores were barely above two-thirds of that figure. The relative deviation was even lower.

The distribution of the most significant indicators is more diverse at farm level than at field level. Of the 17 indicators, seven accounted for an average of 75 per cent of the total score. These indicators are the *diversity of crop structure* (indicator 18; 20 per cent), *frequency of pesticide usage* (indicator 10; 16 per cent), *proportion of cereals and maize* (indicator 16; 11 per cent), *territory of land elements* (indicator 9; 8 per cent), *fuel consumption* (indicator 12; 7 per cent), *nitrogen balance* (indicator 1; 6 per cent) and *non-ploughed areas* (indicator 6; 6 per cent). The results also show that the difference between the relative weights of these highlighted indicators is lower than the same figures at field level.

The data were also analysed according to indicator group. According to the raw data, *nutrient management* (group A) and *natural landscape elements* (group C) had the biggest effects on the variance of the environmental performance of farms (Table 3).

These results are reflected in the distributions of scores across the farms. In the cases of indicators 5, 8, 11 and 13 more than 80 per cent of farms scored zero each year (Table 4), although three of these were simply scored 'yes/no'. This means the farms have the least favourable effect on the environment as regards *winter land coverage, area effected by water management, irrigation* and *external services*. In terms of *non-ploughed areas* and *frequency of pesticide usage* (indicators 6 and 10), more than 10 per cent of the farms scored very highly (meaning that more than 10 per cent of the farms have the most favourable effect in these areas).

Among the four new indicators, *length of field boundary with hedge* (indicator 7) has the highest effect (4.3 per cent on average) on the overall farm-level scores (Table 5).

Table 3: Share of variance by indicator group from raw data and standardised data.

	Indicator group	Raw data	Standardised data
A	Nutrient management (1-4)	71.9	13.5
B	Soil protection (5,6)	6.2	28.1
C	Natural landscape elements (7-9, 16)	13.5	11.0
D	Plant protection (10)	4.1	33.0
E	Energy consumption (11-13)	1.5	10.6
F	Diversity of crop production (14, 15, 17)	2.8	3.8

Source: own data

Table 4: Share of farms achieving the most and least (zero) favourable possible scores per indicator for each individual year (per cent).

Indicator number*	2008/2009		2009/2010		2010/2011	
	least	most	least	most	least	most
1	25	3	22	3	25	3
2	38	4	33	4	37	3
3	53	4	48	4	47	3
4	67	0	65	0	61	0
5	85	5	85	5	85	6
6	23	11	21	10	23	12
7	28	0	28	0	28	0
8	87	1	85	1	87	1
9	28	10	27	9	27	10
10	5	12	6	10	5	10
11	87	2	91	1	89	2
12	8	19	7	18	7	19
13	90	10	90	9	90	10
14	-	-	-	-	-	-
15	52	0	46	0	47	0
16	10	4	10	4	9	5
17	54	24	56	25	55	24
18	237	1	235	1	235	1

* For indicator names see Table 2
Source: own data

Table 5: Proportion of total farm-level scores accounted for by the four new indicators (per cent).

	Length of field boundary with hedge (7)	Area affected by water management (8)	Irrigation (11)	External services (13)
Minimum	0.0	0.0	0.0	0.0
Maximum	15.9	11.1	11.1	5.2
Mean	4.3	0.4	0.4	0.4
Standard deviation	4.2	1.3	1.4	1.2

Source: own data

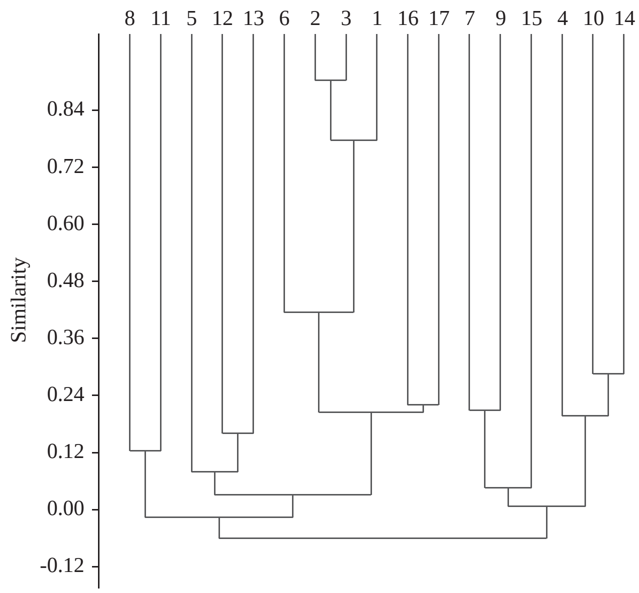


Figure 3: Hierarchical classification of the 17 standardised indicators calculated with cluster analysis (paired group method).
Source: own composition

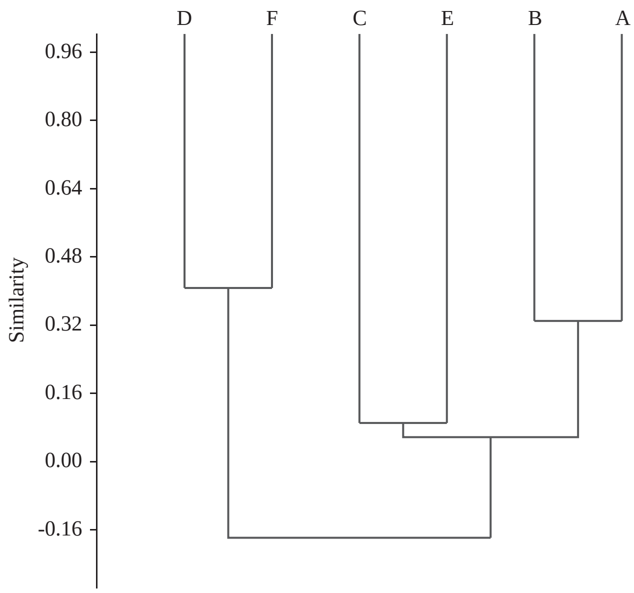


Figure 4: Correlation based cluster analysis of the standardised data of indicator groups A-E.
Source: own composition

Analysis of standardised data

Linear correlation and rank correlation were used on the farm-level data to compare the results of the original calculation method (raw data) and the standardised-balanced summing method (standardised data). Both the linear (0.79) and rank (0.83) correlations were significant and positive, and were significantly different from zero (for linear correlation: p=1.21 *10 ad-57; for rank correlation: p=2.28 *10 ad-67). This means that farms that scored highly using the raw data also scored highly when applying the standardised data.

By applying correlation as a distance function (in similarity form), cluster analysis was used to create a hierarchical classification of the 17 standardised indicators (Figure 3). Zero or negative similarity index values mean that indicators are not similar, whereas a positive value indicates similarity between indicators. Similarity between indicators is shown by horizontal lines. Indicators 1, 2 and 3 are much more strongly related than any other variable pairs and a hiatus is visible in the distribution of correlation values. Indicators 7 and 9, 12-13 and 14-15 also form separate clusters. To overcome these issues, indicators within the six indicator groups A-E were totalled and divided by the number of indicators in the group. These group average values were summed so that the parameters influenced the final indicator values of farms equally. The relationships between the six indicator groups were examined by correlation matrix and correlation-based cluster analysis. All correlations were below 0.45 (Table 6).

Table 6: Correlation matrix of the six indicator groups.

	A	B	C	D	E	F
A	-	5.56E-08	0.010	0.057	0.812	9.45E-14
B	0.329	-	0.864	7.84E-08	0.337	2.89E-10
C	0.160	-0.011	-	0.568	0.152	0.249
D	-0.118	-0.326	0.036	-	0.390	9.65E-12
E	0.015	0.060	0.089	-0.054	-	0.146
F	-0.440	-0.378	-0.072	0.406	-0.090	-

For indicator group names and constituent indicators see Table 2
Source: own calculations

The correlation-based cluster analysis reveals the most similarity between indicator groups A and B (*nutrient management* and *soil protection*) and groups D and F (*plant protection* and *diversity of crop production*). This would seem reasonable as the first two groups are soil-related and the second two are focusing on the produced plants (Figure 4). The one-way ANOVA test shows that there is no significant difference in the average values between the indicator groups B and C, B and E, and C and E, meaning that from these aspects the performance of Hungarian farms are similar. However, all other indicator groups are significantly different from each other. The environmental sustainability the sampled farms with respect to each of the six indicator groups is illustrated in Figure 5. The closer the value is to 1, the higher is the sustainability.

As with the indicator group raw data, the shares of variance in the farm-level scores of the indicator group standardised data were calculated (Table 3). With the elimination of the bias originating from the weighting of the indicators, the variance was influenced mainly by indicator groups A (*nutrient management*), B (*soil protection*) and D (*plant protection*).

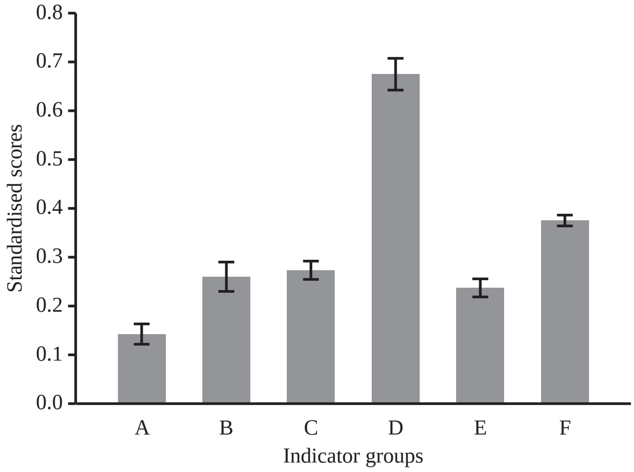


Figure 5: Environmental sustainability of a sample of Hungarian farms according to six indicator groups using standardised data.
Error bars indicate ± 1 SE
Source: own composition

Discussion

The results suggest that the average environmental performance of the farms in our sample is relatively low. This assessment applies both to the field (where the average is 14.1, i.e. 23 per cent of the total available points) and farm (the average of 27.6 is just 33 per cent of the total) scores. Farms scored the highest for *plant protection* and *diversity of crop production*, meaning that their activities are the least environmentally destructive in these areas. The most critical area is *nutrient management*, whereas *soil protection*, *natural landscape elements* and *energy consumption* show average results.

However, the relative weighting of individual indicators can be just as important in determining the apparent level of sustainability as the actual choice of indicators. The Green-point system is still under development and it may be that imperfect weighting is the explanation behind our results. For this reason, the data were standardised to eliminate the influence of weighting. After comparing the standardised and the raw data, differences among the most determinant indicators and indicator groups were outlined. For the raw data, 75 per cent of the variance was determined by two indicator groups (A and C, including six indicators), whereas for the standardised data this level of variance was covered by five indicator groups (D, B, A and C, including eleven indicators). This difference highlighted that standardisation proved to be useful and reconsideration of indicator weightings might be needed.

The standardisation of data is also important for the further development of the tool: while the general analysis of the results allows the comparison of farms; the analysis of standardised data enables researchers to evaluate the system itself. By applying these two approaches in parallel, efficient further development can be achieved.

The testing of the applicability of the new indicators was also essential. The focus areas of these new indicators were selected based on international experiences. Even though similar indicators of different assessment tools are calculated in different ways, the topics they try to evaluate overlap. The new indicator of *irrigation*, for example, is one of the indicators of SMART and it was also added to the new indicator set of DIALECTE. However in the Swiss tool, irrigation similar to the indicator of *water management* is covered by more than one indicator. The indicator *length of field boundary with hedge* is of high importance in the English agri-environmental measures such as Entry Level Stewardship, Organic Entry Level Stewardship and High Level Stewardship, and it proved to be an important indicator of the Green-point system as well as it has a considerable effect on the variance of both the standardised and raw results.

Another important consideration is the scale used for evaluation. Figures presented in Table 4 revealed that the use of a yes/no scale is not recommended in the long run as it can distort the results. However it can be justified during the testing phase of a new indicator, since it still can show the relevance of the topic on which the indicator is focusing.

The necessity of individual indicators was analysed by checking the correlations between indicators. There was a strong and positive correlation between indicators 1, 2 and 3. This shows that it is not necessary to use all three indicators when assessing the nutrient management. It is enough to use only one of them (probably N), as the inclusion of three such similar indicators (with a total of 30 points) can considerably shift the proportion of the weight of indicator groups relative to each other.

To verify our results we compared them with macro-level data. Macro indicators are quite different from micro indicators but still some overlap was found between the Green-point indicators and the sustainable development indicators of KSH (KSH, 2013). KSH applies two macro-level indicators measuring sustainability (nutrient balance for N and P) which are comparable to the Green-point system's indicators. The patterns of our nutrient balance indicators over the period 2008-2011 are in line with the trends of the relevant macro indicators of KSH.

Although we did not analyse specifically the applicability of our tool for different farming systems, it can be stated that the indicators of the Green-point system, in their present form, are mainly suitable for assessing crop production farms. Some of the indicators are not appropriate in the case of specialised farms (e.g. plantations, agro-forestry systems, nurseries, apiaries), therefore such farms are at a disadvantage as regards collecting green points. Animal husbandry farms face similar drawbacks in this respect as indicators for these activities are missing.

Therefore we consider it important for the future to complement the current indicator set with new indicators which enables the Green-point system to assess the environmental performance of all farming systems in Hungary. Such a complete – perceived impact measuring – system could lead to the elaboration of a real impact measuring indicator system, which can expected to be the future basis of the distribution of agricultural subsidies in the CAP.

Acknowledgement

The research was carried out within Szent István University, Faculty of Agriculture and Environmental Sciences (Research Centre of Excellence-9878/2015/FEKUT). The preparation of the article was financially supported by ÖMKi.

References

Bechini, L. and Castoldi, N. (2009): On-farm monitoring of economic and environmental performances of cropping systems: results of a 2-year study at the field scale in northern Italy. Ecological Indicators **9**, 1096-1113. http://dx.doi.org/10.1016/j.ecolind.2008.12.008

Binder, C.R., Feola, G. and Steinberger, J.K. (2010): Considering the normative, systemic and procedural dimensions of indicator based sustainability assessments in agriculture. Environmental Impact Assessment Review **30** (2), 71-81. http://dx.doi.org/10.1016/j.eiar.2009.06.002

Carmona-Torres, C., Parra-López, C., Hinojosa-Rodríguez, A. and Sayadi, S. (2014): Farm-level multifunctionality associated with farming techniques in olive growing: an integrated modeling approach. Agricultural Systems **127**, 97-114. http://dx.doi.org/10.1016/j.agsy.2014.02.001

Degenfelder, L., Lösch, S., Seibert, O. and Groier, M. (2005): Evaluation des mesures agro-environmentales. Annexe 7: etude nationale autriche [Assessment of agro-environmental measures. Annex 7: National study Austria]. Weidenbach-Triesdorf, Germany: Forschungsgruppe Agrar-und Regionalentwicklung Triesdorf.

Fumagalli, M., Acutis, M., Mazzetto, F., Vidotto, F., Sali, G. and Bechini, L. (2011): An analysis of agricultural sustainability of cropping systems in arable and dairy farms in an intensively cultivated plain. European Journal of Agronomy **34**, 71-82. http://dx.doi.org/10.1016/j.eja.2010.11.001

Gras, R., Benoit, M., Deffontaines, J.P., Duru, M., Lafarge, M., Langlet, A. and Osty, P.L. (1989): Le fait technique en agronomie. Activité agricole, concepts et éthodes d'étude [The technique is in agronomy. Agriculture, concepts and methods study]. Paris: L'Harmattan.

Hanley, N., Moffat, I., Faichney, R. and Wilson, M. (1999): Measuring sustainability: A time series of alternative indicators for Scotland. Ecological Economics **28** (1), 55-73. http://dx.doi.org/10.1016/S0921-8009(98)00027-5

Hansen, J.W. (1996): Is agricultural sustainability a useful concept? Agricultural Systems **50**, 117-143. http://dx.doi.org/10.1016/0308-521X(95)00011-S

IEEP (2014): Conference Summary: Results-based agri-environmental schemes: payments for biodiversity achievements in agriculture, Brussel, Belgium, 23-24 September 2014. Available online at http://ec.europa.eu/environment/nature/rbaps/conference/docs/Findings_Conference_Programme-Brussels_23-24_Sept_2014.pdf (accessed 6 December 2014).

Jawtusch, J., Schader, C., Stolze, M., Baumgart, L. and Niggli, U. (2013): Sustainability Monitoring and Assessment Routine: Results from pilot applications of the FAO SAFA Guidelines. Paper presented at the International Symposium on Mediterranean Organic Agriculture and Quality Signs related to the Origin, Agadir, Morocco, 2-4 December 2013.

Kätterer, T. and Andrén, O. (1999): Long-term agricultural field experiments in northern Europe: analysis of the influence of management on soil carbon stocks using the ICBM model. Agriculture, Ecosystems & Environment **72** (2), 165-179. http://dx.doi.org/10.1016/S0167-8809(98)00177-7

KSH (2013): A fenntartható fejlődés indikátorai Magyarországon [The sustainable development indicators in Hungary]. Budapest: KSH.

Meadows, D.H. (1990): A Reaction from a Multitude, in G.M. Woodwell (ed.), The Earth in Transition: Patterns and Processes of Biotic Impoverishment. Cambridge: CUP, 513-521. http://dx.doi.org/10.1017/CBO9780511529917.027

Meul, M., Van Passel, S., Nevens, F., Dessein, J., Rogge, E., Mulier, A. and Van Hauwermeiren, A. (2008): MOTIFS: a monitoring tool for integrated farm sustainability. Agronomy for Sustainable Development **28** (2), 321-332. http://dx.doi.org/10.1051/agro:2008001

Pacini, C., Giesen, G., Wossink, A., Omodei-Zorini, L. and Huirne, R. (2004): The EU's Agenda 2000 reform and the sustainability of organic farming in Tuscany: ecological–economic model at field and farm level. Agricultural Systems **80**, 171-197. http://dx.doi.org/10.1016/j.agsy.2003.07.002

Pimentel, D.C. and Kounang, N. (1998): Ecology of soil erosion in ecosystems. Ecosystems **1**, 416-426. http://dx.doi.org/10.1007/s100219900035

Pointereau, P., Langevin, B. and Gimaret, M. (2012): DIALECTE, a comprehensive and quick tool to assess the agro-environmental performance of farms. Paper presented at the 10th European IFSA Symposium, Aarhus, Denmark, 1-4 July 2012.

Ripoll-Bosch, R., Díez-Unquera, B., Ruiz, R., Villalba, D., Molina, E., Joy, M., Olaizola, A. and Bernués, A. (2012): An integrated sustainability assessment of Mediterranean sheep farms with different degrees of intensification. Agricultural Systems **105**, 46-56. http://dx.doi.org/10.1016/j.agsy.2011.10.003

Ryden, J.C., Ball, P.R. and Garwood, E.A. (1984): Nitrate leaching from grassland. Nature **311**, 50-53. http://dx.doi.org/10.1038/311050a0

Schader, C., Grenz, J., Meier, M.S. and Stolze, M. (2014): Scope and precision of sustainability assessment approaches to food systems. Ecology and Society **19** (3): 42. http://dx.doi.org/10.5751/ES-06866-190342

Singh, R.K., Murty, H.R. Gupta, S.K. and Dikshit, A.K. (2009): An overview of sustainability assessment methodologies. Ecological Indicators **9** (2), 189-212. http://dx.doi.org/10.1016/j.ecolind.2008.05.011

Tangermann, S. (2011): Direct Payments in the CAP post 2013: note. IP/B/AGRI/IC/2011_003. Brussel: European Commission.

Thomassen, M.A., Dolman, M.A., van Calker, K.J. and de Boer, I.J.M. (2009): Relating life cycle assessment indicators to gross value added for Dutch dairy farms. Ecological Economics **68**, 2278-2284. http://dx.doi.org/10.1016/j.ecolecon.2009.02.011

van der Werf, H.M.G. and Petit, J. (2002): Evaluation of the environmental impact of agriculture at the farm level: a comparison and analysis of 12 indicator based methods. Agriculture, Ecosystems & Environment **93** (1-3), 131-145. http://dx.doi.org/10.1016/S0167-8809(01)00354-1

Van Passel, S., Nevens, F., Mathijs, E. and van Huylenbroeck, G. (2007): Measuring farm sustainability and explaining differences in sustainable efficiency. Ecological Economics **62**, 149-161. http://dx.doi.org/10.1016/j.ecolecon.2006.06.008

Wauchope, R.D. (1978): The pesticide content of surface water drainage from agricultural fields. A Review. Journal of Environmental Quality **7** (4), 459-472. http://dx.doi.org/10.2134/jeq1978.00472425000700040001x

Zahm, F., Viaux, P., Girardin, P., Vilain, L. and Mouchet, C. (2006): Farm Sustainability Assessment using the IDEA Method: From the concept of farm sustainability to case studies on French farms, in Proceedings and outputs of the first Symposium of the International Forum on Assessing Sustainability in Agriculture, Bern, Switzerland, 16 March 2006, 77-110.