ALTERNATIVE APPROACHES ON CONSTRUCTING A COMPOSITE INDICATOR TO MEASURE AGRICULTURAL SUSTAINABILITY

José A. Gómez-Limón [a] , Laura Riesgo [b,1]

[a] **Dept. of Agricultural Economics. University of Valladolid. Palencia, Spain**

[b] **Dept. of Economics. Pablo de Olavide University. Seville, Spain**

[1] Corresponding author: laurariesgo@upo.es (Laura Riesgo)

Paper prepared for presentation at the 107th EAAE Seminar "*Modelling of Agricultural and Rural Development Policies*". Sevilla, Spain, January 29th -February 1st, 2008

Abstract

The aim of this paper is to carry out a comparative analysis of alternative methods on constructing composite indicators to measure global sustainability of the agricultural sector. This comparison is implemented empirically on the irrigated agriculture of the Duero basin (Spain) as a case study. For this purpose, this research uses a dataset of indicators previously calculated for different farm-types and policy scenarios. The results allow to establish a hierarchy of the policy scenarios on the basis of the level of sustainability achieved. Furthermore, analyzing the heterogeneity of different farms-types in each scenario, is also possible to determine the main features of the most sustainable farms in each case. All this information is useful in order to support agricultural policy design and its implementation, trying to increase the sustainability of this sector.

Key words: Sustainability, Composite indicators, Irrigated agriculture, Scenarios, Agricultural policy.

1. Introduction

Agricultural sustainability has not a unique meaning. Hansen (1996) identified two broad interpretations of agricultural sustainability. The first one focuses on a normative approach in response to concerns about negative impacts of "conventional" agriculture. This approach relies on the implementation of "alternative" agriculture (ecological agriculture, conservative agriculture, etc.), as an ideological option to achieve a set of values that should characterize this sector. The second meaning follows a positive approach, and it is focused on the ability of agricultural systems to satisfy different demands through time. As it was pointed out by this author, only the latter meaning is useful from a scientific point of view. Thus, in this paper we follow this approach.

However, it is worth pointing out that the selected concept of sustainability has several difficulties to be used empirically in the real world. First, we have to deal with the temporal nature of sustainability. Indeed, this meaning of sustainability related to the preservation of production capacity has little practical value because of the infeasibility of long-term experiments. Second, we have to deal with the difficulty of identifying the demands that must be satisfied by the agricultural sector to be sustainable. In this way, sustainability can be interpreted as a social conception, that can be changed in response to society's requirements. Thus, the meaning of sustainability must be considered local and time specific. Both difficulties have limited for a long time the usefulness of this concept as a criterion for guiding the agricultural development.

In order to avoid the difficulties mentioned above, a wide consensus has been built in order to consider that the sustainability embodies three main dimensions: environmental, economic and social (Yunlong and Smit, 1994). In this sense, it can be assumed that an agricultural system is sustainable when the trade-offs between the objectives considered for public evaluation of its performance (economic objectives –as the income growth or the macroeconomic stability–, social objectives –as the equity or the cover of basic needs–, and ecological objectives –as the ecosystem protection or natural resources regeneration–) reach acceptable values for the society as a whole (Hediger, 1999; Stoorvogel et al., 2004). This approximation to agricultural sustainability makes possible its use as an operational criterion, by using a set of indicators that covers the three dimensions mentioned above.

However, the quantification of the agricultural sustainability through a set of indicators has still some shortcomings. The main inconvenience comes from the difficult interpretation of the whole set of indicators. In order to avoid this problem, it has been suggested that the analysis of agricultural sustainability can be tacked by aggregating this multidimensional set of indicators into a single index or composite indicator. This approach has been used, among others, by Stockle et al. (1994), Andreoli and Tellarini (2000), Pirazzoli and Castellini (2000), Sands and Podmore (2000), Rigby et al. (2001) or van Calker et al. (2006). Nevertheless, the aggregation of indicators has been frequently criticised for: a) the subjectivity of these methods (the choice of functional forms for aggregation and weighting for individual indicators), and b) the compensability usually considered to aggregate the different dimensions or attributes of sustainability (additive aggregation approaches), in spite of their theoretical incommensurability. For further details, the works of Hansen (1996), Bockstaller et al. (1997), Morse et al. (2001), Ebert and Welsch (2004) or Munda (2005) can be consulted.

Within this general framework, this paper has a double objective. First, from a theoretical perspective, this work analyses the pros and cons of alternative methods to build composite indicators to measure agricultural sustainability. This is to be done empirically by implementing these methods to a real world case study. Specifically, we apply these methods to quantify the global sustainability of irrigated agriculture in the Duero basin (inner Spain), using a dataset of indicators previously developed (Riesgo and Gómez-Limón, 2005 and 2006), which covers the three dimensions of sustainability mentioned above. This set of indicators has been calculated for different farm-types and future policy scenarios. This feature of the data has allowed to consider a second objective: to analyse the real possibilities of using the concept of sustainability as a tool for guiding the public management of agriculture. In this sense, the quantitative approach on the basis of the calculation of composite indicators is used: a) to determine a ranking of policy scenarios based of their sustainability, and b) to find out the most sustainable farm-types in each scenario. These results can be useful to public decision making, both from a strategic (encourage policy actions to promote the most sustainable policy scenarios) and tactical (design of higher support to most sustainable farms) points of view.

In order to achieve these objectives, this paper is organized as follows. Section 2 presents the case study, with a detailed description of the indicators dataset used for this research. Section 3 is devoted to an explanation of the methods used to calculate composite indicators for agricultural sustainability. Section 4 presents the results obtained. Finally, Section 5 draws the discussion of the results and the conclusions reached.

2. Case study

2.1. Irrigated agriculture in the Duero basin

The practical application of sustainability needs, first, to determine time and geographical scopes for the analysis. In this paper, the empirical analysis is focused on current irrigated agriculture developed in the Duero basin (inner Spain). This particular agricultural system covers 563,105 hectares, which are mainly devoted to cereal (maize, barley and wheat) and other annual crops (sugar-beet, sunflower or alfalfa). Thus, this is a representative of a continental agricultural system, characterized by extensive farming with low-value-added, low-labour-intensive crops and highly dependent on Common Agricultural Policy (CAP) subsidies.

The interest of this case study is caused by the recent changes in the policy framework design to its public management. First, it is worth noting the recent CAP reform, approved in June 2003 in Luxemburg, that has been implemented in Spain in 2006. Among the novelties introduced by this reform, the most important one is the partial decoupling of public subsidies received by producers. Furthermore, it is worth pointing out the important implications of the approval of the Water Framework Directive (WFD), which force the implementation of a new water pricing policy before 2010, in order to promote sustainable use of water resources. Both normative novelties make the European irrigated agriculture future uncertain. This especially concerns the irrigated sector in the Duero basin, because of its low profitability and high dependence on CAP subsidies. These arguments justify the interest of the analysis about the future sustainability of this case study.

Finally, it is worth mentioning that this work is part of a broader research, developed within the European research project WADI (see Berbel and Gutiérrez, 2006 for a complete reference), where the irrigated agriculture of the Duero basin has been one of the meaningful case studies analysed (Riesgo and Gómez-Limón, 2005 and 2006). In fact, the primary source of the information used in this paper has been obtained in a previous step of this project, as it is described next.

2.2. Simulation of future performance of irrigated agriculture

As most related works in the literature, this research considers the farm as the basic unit for agricultural sustainability analysis (see van der Werf and Petit, 2002). This option has been taken considering the current knowledge of agriculture-ecosystem interactions and the availability of data.

Taken into account the heterogeneity of the irrigated farms in the Duero basin, a detailed typology of these productive units was needed. For this purpose the criteria considered as classifying variables were regarded to farm characteristics (climate, soil and other resources availability) and farmers' features (socio-economic and decision variables). In this way, 22 groups of farms were obtained. A complete description of the features of each farm-type can be found in Riesgo and Gómez-Limón (2006). In any case, the most noticeable point regarding this classification is that resulting farm-types can be considered as representative productive units that can be analysed separately by means of a single simulation model.

Due to the complexity of farmers' decision-making, the programming modelling technique employed to simulate farmers' behaviour facing future policy scenarios has been based on the Multi-Criteria Decision Making (MCDM) paradigm. More concretely, Multi-Attribute Utility Theory (MAUT) has been proposed as a theoretical framework to simulate their decision-making processes. Further details about these models (decision variables, the attributes in the utility function or the constraints) can be seen in Gómez-Limón and Riesgo (2004) and Riesgo and Gómez-Limón (2006). In any case, for this research is only relevant to point out that the simulations developed in this way have allowed to calculate the values of the decision variables and sustainability indicators in each policy scenario.

2.3. Policy scenarios

One of the results of the WADI research project has been the definition of a set of future scenarios of European irrigated agriculture for 2020 (Morris et al., 2004). This project has distinguished four policy scenarios within this time horizon. These scenarios were designed in terms of certain social values (consumerism *vs*. community values) and governance strategies (globalization *vs*. regionalization). The major features of each scenario are as follows:

- *World Markets (WM).* This scenario emphasizes private consumption and a highly developed and integrated world trading system, emphasising economic development (consumerism) and global integration (globalization). In practice, this scenario assumes a fall in agricultural prices due to severe international competition, as well as a rise in yields because of the expansion of genetically modified organisms (GMO). Likewise, this scenario assumes a decrease in the costs of agricultural inputs and the disappearance of agricultural public subsidies, also because of the liberalization of international markets.
- *Global Sustainability (GS).* In this scenario special importance is attributed to social and ecological values (community) in a global economic framework (globalization). Bearing these assumptions in mind, there is collective action to address environmental and social issues. To simulate this scenario, a reduction in agricultural prices due to severe international competition is assumed, although this is less acute than in the first scenario. A rise in crop yields and a moderate reduction in public subsidies in comparison with the current situation are also assumed. Finally, the prices of agricultural inputs will tend to increase, especially those regarded as “pollutants” (fertilizers, pesticides and fossil energy), due to the introduction of eco-taxes.
- *Provincial Enterprise (PE).* This scenario emphasizes private consumption (consumerism), but unlike the other scenarios, it takes account of national or regional level policy decision-making (regionalism), in order to reflect local priorities and interests. To simulate this scenario, it is assumed an increase in agricultural prices because of national agriculture protection policies. On the other hand, agricultural yields will increase due to the introduction of GMO. Finally, agricultural subsidies will be practically the same as the current ones, although the prices of inputs will increase as a result of trade protectionism.
- *Local Stewardship (LS).* In this situation, regional or local governments emphasize social and environmental values (community) that promote local interests (regionalization). This scenario assumes a significant increase in agricultural prices resulting from a higher degree of protectionism. Nevertheless, it should be noted as opposed to the other scenarios, agricultural yields will decrease. This is due to the rejection of GMO and strict controls on the use of inputs. Public subsidies will increase slightly, as will the prices of agricultural inputs as a consequence of environmental taxes charged on the consumption of these products.

Besides these four agricultural policy scenarios, we analysed two additional ones:

- *Statu quo (SQ).* It is considered as the reference scenario since it is the real situation with more recent data available. This scenario describes the situation of the irrigated agricultural sector in 2005, as characterised by CAP defined in the Agenda 2000. The main purpose of this scenario was to enable proper comparisons to be made between the situation in 2005 and the scenarios proposed above.
- *Luxemburg Agreement (LA).* The main purpose of this scenario is to analyse the effects of the implementation of last CAP reform. The main difference with the *statu quo* scenario is the partial decoupling of public subsidies (only 25% of former direct subsidy payments -linked to farmers’

production- are still in force; the other 75% are now paid by a single farm payment fixed for each producer on the basis of a reference period).

For a detailed definition of the above-mentioned agricultural policy scenarios (story lines), the works of Morris et al. (2004) and Riesgo and Gómez-Limón (2006) can be consulted.

2.4. Economic, social and environmental indicators

The research project WADI has done a selection of indicators in order to quantify the sustainability of irrigated agriculture at farm level (see Bazzani et al., 2004, for a detailed reference). This selection is based on the guidelines published by the OECD (OECD, 2001), and it is summarized in Table 1:

Table 1 - Indicators set for the measurement of agricultural sustainability

Area under analysis		**Indicators**	**Measurement units**
Economic sustainability		Total Gross Margin (*TGM*)	€/ha
		Profit (*PROFIT*)	€/ha
		GDP Contribution (*GDPCON*)	€/ha
		Public Subsidies (*PUBSUB*)	€/ha
Social sustainability		Total Labour (*TL*)	person·day/ha
		Seasonal Labour Employment (*SEASONA*)	%
Environmental sustainability	*Landscape and biodiversity*	Agro-diversity (*AGRDIV*)	no. crops
		Soil cover (*SOILCOV*)	%
	Water use	Water use (*WATER*)	m^3/ha
	Fertilizers and pesticides	Nitrogen Balance (*NBAL*)	kg N/ha
		Energy Balance (*EBAL*)	10^6 kcal/ha
		Pesticide Risk (*PESTRISK*)	10^3 RP/ha

Using the simulation technique mentioned before, the value of these sustainability indicators have been calculated for the 22 farm-types representative of the case study in the 6 policy scenarios considered.

3. Methodology

Once the methodology followed to obtain the indicators dataset has been briefly explained, the purpose of this section is to describe the procedures employed to construct a composite indicator to measure agricultural sustainability.

3.1. Methodological framework

The usefulness of the composite indicators has being increasingly recognized to analyse and to communicate complex and multidimensional issues, as it is the case of agricultural sustainability. Such interest has been showed in some recent publications that analyses the alternative methods and techniques to build these types of indices, as are the works of Nardo et al. (2005a and 2005b). These authors have pointed out the different steps that analysts should follow in order to build these composite indicators: 1) Development of the theoretical framework; 2) Basic indicators selection; 3) Multivariate analysis; 4) Imputation of missing data; 5) Normalization; 6) Weighting and aggregation; 7) Robustness and sensitivity; 8) Composite indicators links to other variables; 9) Back to the real data and 10) Presentation and spreading.

The three first steps have been already developed in the Section 2. Step 4 is not necessary in this case study, because the indicators dataset is complete and none imputation of missing data is required.

Normalization (step 5) is a previous requirement to any aggregation of indicators because they are usually measured in different units. In this study, among all existing normalization techniques (Freudenberg, 2003), the one chosen is the re-scaling in a range [0,1]. In this sense, after normalization the scores of indicators range between 0 (the worst value, meaning the least sustainable option) and 1 (the best value, corresponding with the most sustainable option).

Once the normalization is done, indicators should be weighted and aggregated (step 6). Nardo et al. (2005a and 2005b) suggests a number of alternative aggregation techniques. In this work, we select three of these methods, on the basis of a) the principal component analysis, b) the analytic hierarchy process and c) a multi-criteria method founded on the concept of the distance to ideal point. Next sections explain each of these methods used to build the *global sustainability indicator* (GSI) in an operative way. Afterwards, once the different results will be obtained, a critical and comparative analysis of the three methodologies will be done, covering the steps 7, 8 and 9 above mentioned.

3.2. Aggregation method based on Principal Component Analysis (PCA)

A detailed description on constructing indices using Principal Component Analysis (PCA) can be found in Nicoletti et al. (2000) and Nardo et al. (2005a, 2005b). This section summarizes the application of this technique to our case study, focused on the construction of the GSI-PCA.

In this work, the PCA technique is applied to the indicators dataset describing the *statu quo* (SQ) scenario (22 farm-types × 12 indicators), in order to group those indicators highly correlated. In this way, the principal components Z_j are obtained. For this purpose, only those principal components with eigenvalues higher than one are retained. Furthermore, to facilitate the interpretation of these components a Kaiser's varimax rotation is implemented. The results obtained can be observed in the Table 2.

Table 2 – Principal components extracted to build the GSI-PCA

Components	**Extraction sum of squared loadings**			**Rotation sum of squared loadings**		
	Total (eigenvalue)	*% of variance*	*Cumulative %*	*Total (eigenvalue)*	*% of variance*	*Cumulative %*
Z_1	6.733	56.110	56.110	4.615	38.458	38.458
Z_2	1.990	16.587	72.696	3.524	29.363	67.821
Z_3	1.635	13.624	86.320	2.220	18.499	86.320
Kaiser-Meyer-Olkin measure of sampling adequacy					0.594	
Bartlett's test of sphericity				$\chi^2 = 363.26$	g.l. = 66	$p < 0.0001$

Taking into account these results, 3 principal components are retained, explaining the 86.3% of the total variance. To understand the meaning of these components, rotated factor loadings of the different indicators can be analysed, as it can be seen in Table 3.

Table 3 - Rotated components matrix from PCA (factor loadings)

Indicators	**Z_1**	**Z_2**	**Z_3**	**Communalities**
TGM	0.605	**0.685**	0.364	*0.967*
PROFIT	-0.015	**0.769**	0.393	*0.746*
GDPCON	**0.727**	0.419	0.454	*0.910*
PUBSUB	0.244	**0.925**	-0.059	*0.919*
TL	**0.800**	0.383	0.240	*0.845*
SEASONA	-0.220	**-0.811**	0.298	*0.795*
AGRDIV	-0.339	-0.335	**-0.766**	*0.813*
SOILCOV	-0.010	-0.190	**0.924**	*0.890*
WATER	**-0.908**	-0.164	0.085	*0.858*
NBAL	**-0.666**	-0.642	-0.126	*0.871*
EBAL	**-0.885**	-0.193	0.099	*0.830*
PESTRISK	**0.898**	-0.051	0.324	*0.914*

Once these principal components are extracted, the calculation of intermediate sustainability indicators (ISI_j), corresponding to each of the principal component *j*, are needed. This is done by calculating a weighted aggregation of indicators:

$$ISI_{ji} = \sum_{k=1}^{k=n} w_{kj} I_{ki} \qquad [1]$$

where ISI_{ji} is the intermediate sustainability indicator for the component *j* and the farm *i*, w_{kj} represents the weight of indicator *k* in the component *j* and I_{ki} is the normalized indicator *k* achieved by the farm *i*.

The weights w_{kj} are obtained from the factor loadings matrix above mentioned following this expression:

$$w_{kj} = \frac{\left(factor\_loading_{kj}\right)^2}{eigenvalue_j} \qquad [2]$$

where $factor\_loading_{kj}$ is the value of the factor loading of indicator *k* in the principal component *j* (see Table 3), and $eigenvalue_j$ is the eigenvalue of the *j*th principal component (see Table 2).

Finally, the GSI-PCA can be calculated as a weighted aggregation of the intermediate sustainability indicators:

$$GSI - PCA_i = \sum_{j=1}^{j=3} \alpha_j ISI_{ji} \qquad [3]$$

where $GSI\text{-}PCA_i$ is the value of the composite indicator for the farm *i* and α_j is the weight applied to the intermediate sustainability indicator *j*. These weights are calculates as follows:

$$\alpha_j = \frac{eigenvalue_j}{\sum_{j=1}^{j=3} eigenvalue_j} \qquad [4]$$

3.3. Aggregation method based on Analytic Hierarchy Process (AHP)

The second approach selected for building the GSI is the Analytic Hierarchy Process (AHP). This technique has been widely adopted as a means of making complex decisions, but it can be adapted for constructing composite indicators (see Nardo et al., 2005a, 2005b).

The AHP method was created by Saaty (1980) as a structured but flexible technique for making decisions in a multi-criteria context. This method is based on approaching complex decision problems using a hierarchical structure. Figure 1 shows the four-level structure considered for our case study.

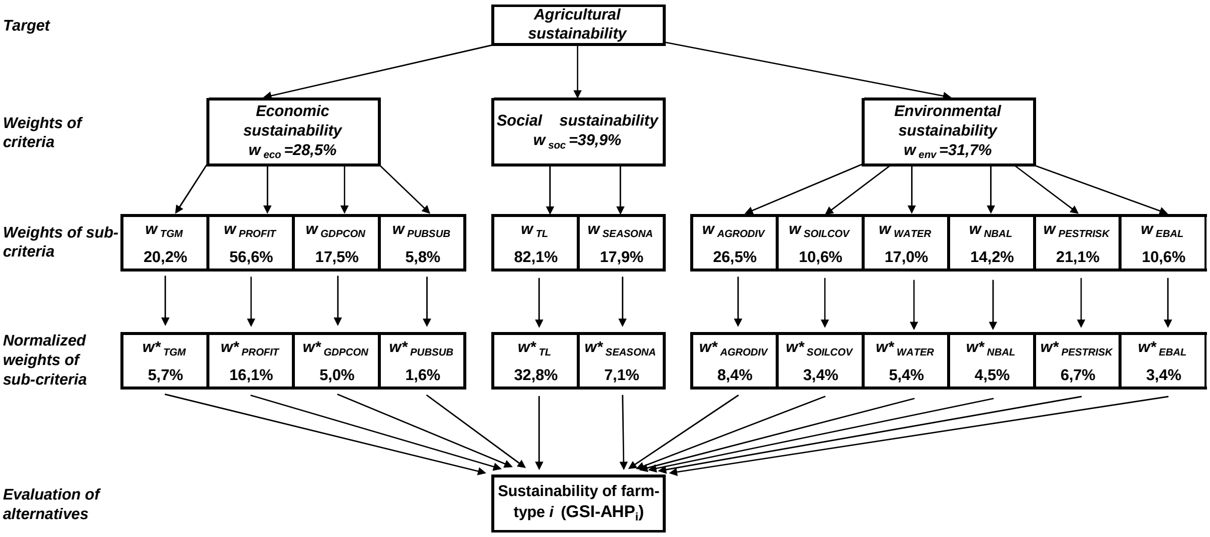

Figure 1 - Hierarchical structure to construct the GSI-AHP

Within this hierarchical structure, the relative importance or weightings (w_k) of criteria or sub-criteria hanging on each node are obtained from pair-wise comparisons between them. In order to perform these pair-wise comparisons, a 1-9 scale is used, as has been proposed by Saaty (1980). Scores of these comparisons are used to build the Saaty's matrices ($A = a_{kl}$), which are employed in order to determine the vector of priorities or weights ($w_1,\dots w_k,\dots w_n$). Although different procedures to estimate these weights have been proposed, for this case we select the simplest one: the geometric mean method (Aguarón and Moreno-Jiménez, 2000).

Initially, the AHP decision technique was designed for individual decision-makers, but was promptly extended for group decisions (Easley et al., 2000). The latter is our case study. Thus to determine the weights attached to each criterion we have to consider the judgements of a group of people (p), each one with his/her own pair-wise comparison matrix ($A_p = a_{klp}$) and the related weights (w_{kp}). This individual information is properly treated in order to obtain a synthesis of aggregated weights (w_k). For this purpose, Forman and Peniwati (1998) suggest that group decision making should be done aggregating individual priorities using the geometric mean:

$$w_k = \sqrt[m]{\prod_{p=1}^{p=m} w_{kp}} \qquad [5]$$

At this point it is worth identifying the most appropriate respondents to assign the weights (w_k) required on the build the GSI-AHP. For this research is considered that the relative importance of each criterion (importance of economic, social and environmental sustainability within the global sustainability) should be assigned by the whole society. In this sense, as was pointed out in the introduction section, it is assumed that sustainability is a social conception. Opposite, for the sub-criteria weighting (importance of indicators within the economic, social and environmental sustainability) it is more adequate to consider the opinion of experts on agricultural sustainability, due to the technical nature of the comparison to be done.

Regarding the weighting of the criteria, it is worth mentioning the work developed by Gómez-Limón and Atance (2004). These authors addressed the relative importance of public objectives that should guide the agricultural policy in the region where the Duero basin is located (*Castilla y León*). To achieve this objective they applied the AHP method to a hierarchical structure where the criteria were identified with economic, social and environmental objectives. In order to estimate the weights assigned to each generic objective they developed a survey of citizens, obtaining a sample of 321 valid questionnaires (pair comparisons to build individual Saaty's matrices). Because of the similarity between the generic objectives considered in this study and the three basic dimensions of sustainability analysed in this paper, we judge suitable the use of the results obtained by Gómez-Limón and Atance (2004) for weighting our criteria. Thus, we have: w_{eco}=28.5%, w_{soc}=39.9% and w_{env}=31.7% (see second row in Figure 1).

For the case of sub-criteria (indicators) weighting, we employ a panel of 10 experts in agricultural sustainability (university lecturers, members from agricultural research centres and civil servants in charge of agricultural policy implementation). Taking into account the technical information (pair-wise comparison) provided by the panel of experts, and following the procedure explained above, the relative weights for the different indicators were obtained at aggregated level (w_k), as it can be seen in Figure 1. In any case, to make operative such weights, it is necessary to normalise them (normalised weights w_k^* should add one). To fulfil these requirements the weight of each sub-criterion (indicator) is multiplied by the weight of its own criterion (importance of economic, social or environmental sustainability). The final results can be seen in the forth row in Figure 1.

Once the normalised weights are obtained, it is worth mentioning that resolving problems by means of the AHP technique is equivalent to optimising a multi-attribute utility function, as has been proved by Zahedi (1987). Adjusting this formulation to our case study, the GSI-AHP can be obtained through the following expression:

$$GSI-AHP_i = \sum_{k=1}^{k=n} w_k^* \cdot I_{ki} \qquad [6]$$

where $GSI\text{-}AHP_i$ is the global sustainability by farm i, w_k^* is the normalised weight to indicator k, and I_{ki} is the normalised outcome of indicator k in the farm i.

3.4. Aggregation method based on a multi-criteria technique (MCDM)

The third method followed in this paper to calculate the GSI is based on multi-criteria techniques (GSI-MCDM), as has been developed by Díaz-Balteiro and Romero (2004). Using this approach, the construction of a sustainability index can be done by calculating the distance between the observed outcomes of the whole set of sustainability indicators and the "ideal point", defined as the hypothetical situation where the values of all indicators reach their most sustainable values. That is, it is possible to quantify the global sustainability of the farm *i* using the following expression:

$$GSI - MCDM_i = \sum_{k=1}^{k=n} w_k^p \left(I_k^* - I_{ki} \right)^p \qquad [7]$$

where w_k is the weight to indicator *k*, I^*_k is the normalised ideal value of indicator *k*, I_{ki} is the normalised value of indicator *k* achieved by farm *i*, and *p* is the metric used to quantify the distances. Thus, the most sustainable farm-type is the one that minimises the value of this index.

Considering that I^*_k=(1,…,1), expression [7] can be modified in order to obtain this new index:

$$GSI - MCDM_i = \sum_{k=1}^{k=n} w_k^p \cdot I_{ki}^p \qquad [8]$$

Now, the most sustainable farm-type is the one that maximises the value of this index.

In any case, both [7] and [8] are generic expressions that depend on the value of the metric *p*. Hence, different values of GSI-MCDM can be obtained for each value of *p*. Thus, for p=1, the global sustainability of farm *i* can be calculated as a weighted sum of the normalised indicators:

$$GSI - MCDM_i = \sum_{k=1}^{k=n} w_k \cdot I_{ki} \qquad [9]$$

As it can be seen, the expression [9] is exactly the same as [6] used for the calculation of GSI-AHP. For this reason, the methodological approach mentioned in the previous section can be considered as a particular case of this more general approach. In this case (p=1), the GSI-MCDM index allows perfect compensability between the different indicators of sustainability (total commensurability of the indicators).

Opposite, when p=∞, the GSI-MCDM quantifies the minimum weighted and normalised value for the set of indicators:

$$GSI - MCDM_i = \underset{k}{Min}\left(w_k \cdot I_{ki} \right) \qquad [10]$$

Thus, the most sustainable farm is the one with the most balanced performance, as long as none of its indicators has values far away from its ideal scores. In this sense, it should be noted that in this particular case (p=∞), against to p=1, no compensation between the indicators is allowed. Therefore, total incommensurability of indicators is assumed.

Of course, any other intermediate metrics (values of *p* within the range [0, ∞]) are possible, assuming different degrees of commensurability between the sustainability indicators. As it is pointed out by Díaz-Balteiro and Romero (2004), a way to trace out all the existing composite indicators from p=1 to p=∞ is to calculate a convex combination of the measures of sustainability given by the expressions [9] and [10]:

$$GSI - MCDM_i = (1-\lambda) \cdot \left[\underset{k}{Min}\left(w_k \cdot I_{ki} \right) \right] + \lambda \cdot \sum_{k=1}^{k=n} w_k \cdot I_{ki} \qquad [11]$$

where λ is a parameter, bounded between 0 and 1, that measures the level of incommensurability between the individual indicators used to calculate the GSI-MCDM. For λ=1 the global sustainability corresponds to those given by the expression [9], considering a total compensability between indicators (AHP approach). For λ=0, non-compensability between indicators is assumed, and therefore the global sustainability corresponds to those given by the expression [10]. For 0<λ<1 partial incommensurability between the indicators considered for the GSI-MCDM is accepted.

To illustrate this methodology we select 5 values for the parameter of incommensurability: λ=0, λ=0.25, λ=0.5, λ=0.75 and λ=1. Moreover, we consider the same weights (w_k) previously calculated for the AHP approach.

4. Results

Tables in the Annex show the results of the GSI obtained following the three aggregation methods described in the previous section. These results are shown for each farm-type and policy scenario.

4.1. Aggregation method based on PCA

Taking into account the results of the GSI-PCA, a first analysis is proposed in order to check if a ranking of policy scenarios can be established on the basis of the degree of sustainability provided. For this purpose, pair comparisons between policy scenarios are done using paired sample *t*-tests for equality of means. Table 4 shows the results of these tests.

Table 4 - GSI-PCA: *t*-tests for global comparison between policy scenarios

Scenario comparison	**Paired differences**			***t***	**d.f.**	***p*-value**
	Mean	*Std. deviation*	*Std. error mean*			
SQ – LA	-0.03655	0.07333	0.01563	-2.338	21	0.029
SQ – WM	0.05795	0.06933	0.01478	3.921	21	0.001
SQ – GS	0.03109	0.06527	0.01392	2.234	21	0.036
SQ – PE	0.02527	0.07429	0.01584	1.596	21	0.126
SQ – LS	0.02468	0.06514	0.01389	1.777	21	0.090
LA – WM	0.09450	0.09069	0.01933	4.888	21	0.000
LA – GS	0.06764	0.08662	0.01847	3.662	21	0.001
LA – PE	0.06182	0.08810	0.01878	3.291	21	0.003
LA – LS	0.06123	0.07123	0.01519	4.032	21	0.001
WM – GS	-0.02686	0.05056	0.01078	-2.492	21	0.021
WM – PE	-0.03268	0.05870	0.01252	-2.611	21	0.016
WM – LS	-0.03327	0.04546	0.00969	-3.433	21	0.002
GS – PE	-0.00582	0.04945	0.01054	-0.552	21	0.587
GS – LS	-0.00641	0.04526	0.00965	-0.664	21	0.514
PE – LS	-0.00059	0.06177	0.01317	-0.045	21	0.965

Results show significant differences (*p*-value<0.05) between most of pairs of scenarios compared. Taking into account these results, the most sustainable scenario is the one established by the

Luxemburg Agreement (LA), following by the policy scenario in force up to 2006 (SQ). Opposite, the least sustainable scenario is the world markets (WM).

From these results, as a policy recommendation to improve the sustainability of the irrigated agriculture in the Duero, it is inferred the convenience of maintaining a moderate interventionism in this sector, preferentially based on the partial decoupling of public subsidies. Thus, these results support the last reform of the CAP (Luxemburg agreement) as a policy change that promotes the improvement of the agricultural sustainability in the case study considered. Moreover, it is worth mentioning that a further liberalization of international markets and/or the withdrawal of the public subsidies to this sector would imply a decrease in the global sustainability of this particular agricultural system.

The second analysis performed is focused on the heterogeneity of farm-types regarding their degree of sustainability in different policy scenarios. For this purpose, a cluster analysis is done in order to classify the farm-types studied into homogeneous groups regarding its level of sustainability. Thus, to develop this classification the values of GSI-PCA$_i$ are considered as classifying variables. Furthermore, a hierarchical aggregation method is applied (the Ward or minimum distance method), defining the distance between elements as the Euclidean square distance. This aggregation procedure results in a dendrogram, that should be "cut" by the researcher to decide the number of clusters considered.

In this particular case, we found appropriate to cut the dendrogram in order to group the 22 fam-types into two homogeneous groups or clusters. Taking into account socio-demographic and decision variables of each of the farms, these groups can be characterized as follows:

- *Cluster PCA-1*. This group comprises small-medium size farms (26.8 ha on average) whose owners are engaged full-time in agriculture (more than 90% of their incomes come from farming). These farms are devoted to the most profitable and value-added crops in this area, such as maize (64%), sugar-beet (10%) and vegetables (4%). These features led us to denominate this group as "*small-medium full-time commercial farmers*".
- *Cluster PCA-2*. This group includes large-medium farms (40.5 ha on average) whose holders were engaged only at part time in farming (more than 20% of their incomes are not coming form agricultural activities). These farms mainly sow winter cereals (36%), maize (31%) and alfalfa (11%). This crop plan is less profitable than the one of cluster PCA-1, but less risky. These features regard this group as conservative farmers that prefer a security income from winter cereals instead of risky crops with higher profit expectancy. Thus, this group is labelled as "*large part-time conservative farmers*".

Once this classification of farms is established, we analyze the differences in the global sustainability (GSI-PCA) of both groups in each of the policy scenarios considered. We use the independent sample *t*-test for equality of means. Table 5 shows the results obtained.

Table 5 - GSI-PCA: *t*-tests for comparison between farm groups

Scenario	**Group Statistics**				***t*-test for equality of means (equal variances not assumed)**			
	Cluster	*N*	*Mean*	*Std. deviat.*	*Mean differences*	*t*	*d.f.*	*p-value*
SQ	*PCA-1*	7	0.5821	0.02867	0.10901	5.880	19.835	0.000
	PCA -2	15	0.4731	0.05825				
LA	*PCA -1*	7	0.6223	0.03045	0.11429	4.607	19.339	0.000
	PCA -2	15	0.5080	0.08511				
WM	*PCA -1*	7	0.5133	0.03392	0.09302	4.634	18.950	0.000
	PCA -2	15	0.4203	0.05982				
GS	*PCA -1*	7	0.5390	0.01472	0.09133	7.371	19.119	0.000
	PCA -2	15	0.4477	0.04288				
PE	*PCA -1*	7	0.5049	0.03777	0.03272	1.759	14.304	0.100
	PCA -2	15	0.4721	0.04617				
LS	*PCA -1*	7	0.5347	0.02824	0.07565	5.242	15.500	0.000
	PCA -2	15	0.4591	0.03762				

Results show that farms belonging to cluster PCA-1 are significantly more sustainable than those include in the cluster PCA-2 for all policy scenarios. Thus, the global sustainability of irrigated agriculture in the Duero basin also depends on the profile of its farms. In this sense, we can conclude that any agricultural policy which main objective was the improvement of the agricultural sustainability should reinforce those farms featured as the cluster PCA-1 (i.e., higher levels of public support).

4.2. Aggregation method based on AHP

To analyse the results for the GSI-AHP we follow the same outline as mentioned above for the GSI-PCA. Thus, first, we explore if a hierarchy of policy scenarios depending on their degree of sustainability exists. For this purpose paired sample *t*-tests for equality of means are applied. Table 6 shows the results of these tests.

Table 6 - GSI-AHP: *t*-tests for global comparison between policy scenarios

Scenario comparison	**Paired differences**			***t***	**d.f.**	***p*-value**
	Mean	*Std. deviation*	*Std. error mean*			
SQ - LA	0.0145	0.0848	0.0181	0.800	21	0.433
SQ - WM	0.0384	0.0974	0.0208	1.848	21	0.079
SQ – GS	0.0252	0.0840	0.0179	1.407	21	0.174
SQ – PE	0.0275	0.0995	0.0212	1.296	21	0.209
SQ – LS	0.0348	0.0878	0.0187	1.860	21	0.077
LA – WM	0.0239	0.1017	0.0217	1.103	21	0.283
LA – GS	0.0107	0.0898	0.0191	0.560	21	0.581
LA – PE	0.0130	0.1168	0.0249	0.524	21	0.606
LA – LS	0.0204	0.0755	0.0161	1.265	21	0.220
WM – GS	-0.0132	0.0752	0.0160	-0.822	21	0.420
WM – PE	-0.0109	0.0998	0.0213	-0.511	21	0.615
WM – LS	-0.0035	0.0660	0.0141	-0.252	21	0.804
GS – PE	0.0023	0.0821	0.0175	0.132	21	0.896
GS – LS	0.0096	0.0501	0.0107	0.902	21	0.377
PE – LS	0.0073	0.0941	0.0201	0.365	21	0.719

In this case we cannot distinguish significant differences (p-value<0.05) for any pair of the scenarios. Therefore, a ranking for the policy scenarios cannot be determined; all scenarios show a similar global sustainability measured as GSI-AHP.

However, if we analyse the partial scores of the different dimensions of sustainability, we can find significant differences among policy scenarios. In this sense it can be pointed out that the economic sustainability in GS scenario is significantly lower than in the rest of the scenarios, or that the environmental sustainability in SQ, LA and GS scenarios is higher than in WM, PE and LS. In this way, it should be noted that the additivity of the AHP method makes that changes in the different dimensions of sustainability in each scenario were compensated each other. This causes that no significant differences in sustainability can be found out at aggregated level, as measured by the GSI-AHP.

The second analysis regarding GSI-AHP results evaluates the heterogeneity of farm-types. For this purpose a cluster analysis is done, following the same technical options than in the case of GSI-PCA. Taking into account the resulting dendrogram, two groups of farms are also considered. Analysing the features of both clusters, we find a high degree of similarity with those obtained for the GSI-PCA. Thus, we can characterize one group as "*large part-time conservative farmers*" (cluster AHP-1, equivalent to PCA-2) and the other one as "*small-medium full-time commercial farmers*" (cluster AHP-2, equivalent to PCA-1).

Once the farm-types are classified into groups, next we explore the differences between these clusters regarding their global sustainability. We also use the independent sample *t*-tests for equality of means. Table 7 shows the results obtained in this way.

Table 7 - GSI-AHP: *t*-tests for comparison between farm groups

Scenario	**Group Statistics**				***t*-test for equality of means (equal variances not assumed)**			
	Cluster	*N*	*Mean*	*Std. deviat.*	*Mean differences*	*t*	*d.f.*	*p-value*
SQ	*AHP-1*	13	0.4464	0.0720	0.1575	7.071	17.068	0.000
	AHP-2	9	0.6039	0.0297				
LA	*AHP-1*	13	0.4462	0.0798	0.1227	3.609	17.690	0.002
	AHP-2	9	0.5689	0.0775				
WM	*AHP-1*	13	0.4067	0.0767	0.1608	6.402	18.894	0.000
	AHP-2	9	0.5674	0.0400				
GS	*AHP-1*	13	0.4188	0.0783	0.1633	6.809	16.638	0.000
	AHP-2	9	0.5821	0.0305				
PE	*AHP-1*	13	0.4501	0.0905	0.0813	2.484	19.988	0.022
	AHP-2	9	0.5313	0.0630				
LS	*AHP-1*	13	0.4242	0.0599	0.1267	5.966	19.991	0.000
	AHP-2	9	0.5509	0.0398				

The results are similar to those obtained for the GSI-PCA: the farms included in the cluster AHP-2 are more sustainable than the ones integrated in the cluster AHP-1 for all policy scenarios. Therefore, conclusions derived from this analysis to public decision-makers are the same than the ones drawn in Section 4.1.

4.3. Aggregation method based on a multi-criteria technique (MCDM)

The different values of the incommensurability parameter (λ) considered creates 5 different expressions for the GSI-MCDM. For the particular case of λ=1, the results obtained are the same than those for GSI-AHP. Thus, the analysis of these results can be also consulted in the previous section. In order to summarize the description of the other results, we only present those obtained for λ=0.5 and λ=0. In this way we try to show the effects of partial (λ=0.5) and total (λ=0) incommensurability on the scores of GSI-MCDM. In any case, the results of GSI-MCDM for λ=0.75 and λ=0.25 will be also commented when necessary to draw more general conclusions.

To analyse the results of GSI-MCDM for λ=0.5 and λ=0, the same procedures as in previous sections are also followed. Table 8 shows the results of the *t*-tests implemented in order to determine if a hierarchy of policy scenarios regarding their sustainability really exits.

Table 8 - GSI-MCDM for λ=0.5 and λ=0: *t*-tests for global comparison between policy scenarios

Value λ	**Scenario comparison**	**Paired differences**			***t***	**d.f.**	***p*-value**
		Mean	*Std. deviation*	*Std. error mean*			
λ=0.5	SQ – LA	0.0067	0.0428	0.0091	0.735	21	0.471
	SQ – WM	0.0207	0.0494	0.0105	1.969	21	0.062
	SQ – GS	0.0129	0.0425	0.0906	1.428	21	0.168
	SQ – PE	0.0126	0.0511	0.0109	1.161	21	0.259
	SQ – LS	0.0172	0.0445	0.0095	1.819	21	0.083
	LA – WM	0.0140	0.0515	0.0110	1.277	21	0.216
	LA – GS	0.0062	0.0451	0.0096	0.646	21	0.525
	LA – PE	0.0059	0.0594	0.0127	0.468	21	0.644
	LA – LS	0.0105	0.0388	0.0083	1.272	21	0.217
	WM – GS	-0.0078	0.0380	0.0081	-0.962	21	0.347
	WM – PE	-0.0081	0.0503	0.0107	-0.754	21	0.459
	WM – LS	-0.0035	0.0331	0.0070	-0.495	21	0.626
	GS – PE	-0.0003	0.0419	0.0089	-0.032	21	0.975
	GS – LS	0.0043	0.0247	0.0053	0.816	21	0.423
	PE – LS	0.0046	0.0480	0.0102	0.449	21	0.658
λ=0	SQ – LA	-0.0011	0.0035	0.0008	-1.456	21	0.160
	SQ – WM	0.0029	0.0042	0.0009	3.273	21	0.004
	SQ – GS	0.0006	0.0040	0.0009	0.754	21	0.459
	SQ – PE	-0.0023	0.0037	0.0008	-2.875	21	0.009
	SQ – LS	-0.0004	0.0036	0.0008	-0.550	21	0.588
	LA – WM	0.0040	0.0040	0.0009	4.659	21	0.000
	LA – GS	0.0017	0.0036	0.0008	2.271	21	0.034
	LA – PE	-0.0115	0.0047	0.0010	-1.154	21	0.261
	LA – LS	0.0007	0.0038	0.0008	0.833	21	0.414
	WM – GS	-0.0023	0.0030	0.0006	-3.511	21	0.002
	WM – PE	-0.0051	0.0034	0.0007	-7.023	21	0.000
	WM – LS	-0.0033	0.0034	0.0007	-4.546	21	0.000
	GS – PE	-0.0029	0.0035	0.0008	-3.833	21	0.001
	GS – LS	-0.0011	0.0033	0.0007	-1.499	21	0.149
	PE – LS	0.0018	0.0038	0.0008	2.268	21	0.034

As it can be observed, the results of this analysis vary widely depending on the value of the incommensurability parameter. In the case where partial compensability (λ=0.5) between the indicators is assumed, the results do not show any significant differences (p-value<0.05) for the different policy scenarios comparison performed. Thus, we cannot establish any ranking of scenarios in this case; all scenarios reach a similar degree of global sustainability in terms of GSI-MCDM for λ=0.5. In this sense it is also worth noting that the same results are obtained for λ=0.25, λ=0.75 and λ=1. For all these values of λ, the results are similar to those explained in Section 4.2 for GSI-AHP.

Opposite, if non-compensability between indicators is assumed (λ=0), it can be observed significant differences (p-value<0.05) for most of the comparison of scenarios implemented. In this case, the most sustainable policy scenarios are Luxemburg agreement (LA) and provincial enterprise (PE), and the least sustainable one is world markets (WM). Taking into account these results, the most sustainable agricultural policies should be characterized by a moderate public interventionism. Moreover, a further liberalization of agricultural markets and a decreasing in public subsidies would lead to a lower sustainability for the irrigated agriculture in the Duero basin. These results are equivalent to those obtained for the GSI-PCA (see Section 4.1).

Using the GSI-MCDM, the basic dimensions of the sustainability (economic, social and environmental) can be also measured, as it was done for the GSI-AHP. In this sense it can be pointed out that for all values of the incommensurability parameter, the lowest economic sustainability is located in GS scenario, while the highest environmental sustainability is reached in SQ and LA scenarios. Furthermore, in this way it should be noted that for those values of λ that allow partial or total compensability between indicators (λ>0), changes in the partial scores of the different dimensions of sustainability are compensated each other. This circumstance makes the global sustainability measured as GSI-MCDM were similar in all the policy scenario considered, as in the case of the GSI-AHP (see Section 4.1).

The second analysis regarding GSI-MCDM results is focused on the heterogeneity of farm-types. For this purpose, a cluster analysis is implemented as in the previous cases. Looking at the dendrograms obtained for the GSI-MCDM for λ=0.5 y λ=0, we decided to classify the 22 farms into 2 groups in both cases.

In the case where partial commensurability is assumed (λ=0.5), the features of the homogeneous groups are the same as those commented for the GSI-AHP (λ=1). In fact, the results of the cluster analysis for λ=1, λ=0.75 and λ=0.5 are exactly the same, and for λ=0.25 the clusters obtained are also very similar. Therefore, for all these cases we can distinguish a group of "*large part-time conservative farmers*" (cluster MCDMa-1, equivalent to AHP-2) and a group of "*small-medium full-time commercial farmers*" (cluster MCDMa-2, equivalent to AHP-2).

However, in the case where λ=0 cluster analysis yields a slightly different typology. In this case, the first group (cluster MCDMb-1) includes farmers involved at part-time in agricultural activities (more than 20% of their incomes depends on non-farming activities), which crop plan comprises mainly maize (42%) and winter cereals (29%). These features led us to denominate this group as "*part-time risk diversified farmers*". The second group (cluster MCDMb-2) is integrated by full-time farmers that crop mainly maize (61%) and other high profitable crops as sugar-beet or vegetables. We labelled this group as "*full-time commercial farmers*". As it can be seen, in this classification the size of the farms is not a significant variable to distinguish between clusters.

Once more, the last analysis tries to find differences in the global sustainability between groups of farms in each policy scenario. For this purpose, *t*-tests for equality of means are applied again. Table 9 shows the results in the case where partial commensurability between indicators is assumed (λ=0.5).

Table 9 - GSI-MCDM for λ=0.5: *t*-tests for comparison between farm groups

Scenario	**Group Statistics**				***t*-test for equality of means (equal variances not assumed)**			
	Cluster	*N*	*Mean*	*Std. deviat.*	*Mean differences*	*t*	*d.f.*	*p-value*
SQ	*MCDMa-1*	13	0.2241	0.0167	0.0803	6.921	17.827	0.000
	MCDMa-2	9	0.3044	0.0368				
LA	*MCDMa-1*	13	0.2245	0.0408	0.0629	3.598	17.602	0.002
	MCDMa-2	9	0.2874	0.0399				
WM	*MCDMa-1*	13	0.2033	0.0383	0.0803	6.407	18.888	0.000
	MCDMa-2	9	0.2837	0.0199				
GS	*MCDMa-1*	13	0.2103	0.0391	0.0823	6.735	17.414	0.000
	MCDMa-2	9	0.2926	0.0168				
PE	*MCDMa-1*	13	0.2278	0.0463	0.0403	2.462	19.975	0.023
	MCDMa-2	9	0.2680	0.0304				
LS	*MCDMa-1*	13	0.2132	0.0302	0.0647	6.040	19.989	0.000
	MCDMa-2	9	0.2779	0.0200				

In this case, the results are the same to those obtained for the GSI-PCA and the GSI-AHP. Thus, the farms included in cluster MCDMa-2 ("*small-medium full-time commercial farmers*") are more sustainable than the ones belonging to cluster MCDMa-1 ("*large part-time conservative farmers*") for all policy scenarios. Then, the conclusions derived from these results are identical that those drawn in previous sections: the public management of the irrigated agriculture in the Duero basin should promote farms with similar profile than described for cluster MCDMa-2.

Table 10 shows the results for the GSI-MCDM for λ=0 (non-commensurability between the indicators is assumed).

Table 10 - GSI-MCDM for λ=0: *t*-tests for comparison between farm groups

Scenario	**Group Statistics**				***t*-test for equality of means (equal variances not assumed)**			
	Cluster	*N*	*Mean*	*Std. deviat.*	*Mean differences*	*t*	*d.f.*	*p-value*
SQ	*MCDMb-1*	14	0.0006	0.0011	-0.0065	-4.075	7.516	0.004
	MCDMb-2	8	0.0070	0.0044				
LA	*MCDMb-1*	14	0.0014	0.0017	-0.0071	-6.998	10.771	0.000
	MCDMb-2	8	0.0085	0.0025				
WM*	*MCDMb-1*	14	0.0000	0.0000				
	MCDMb-2	8	0.0000	0.0000				
GS	*MCDMb-1*	14	0.0011	0.0017	-0.0032	-2.270	8.743	0.050
	MCDMb-2	8	0.0043	0.0038				
PE	*MCDMb-1*	14	0.0044	0.0036	-0.0021	-1.544	17.787	0.140
	MCDMb-2	8	0.0065	0.0028				
LS	*MCDMb-1*	14	0.0020	0.0028	-0.0036	-2.537	12.277	0.026
	MCDMb-2	8	0.0056	0.0034				

* *t* cannot be computed because the standard deviations of both groups are cero.

These results show that farms belonging to the group "*full-time commercial farmers*" (cluster MCDMb-2) are more sustainable than those managed by "*partial-time risk-diversified farmers*" (cluster MCDMb-1) for most of the policy scenarios considered (all of them except scenarios WM and PE). In this sense, we can conclude that any agricultural policy which pursues the promotion of sustainability in this agricultural system should encourage the reinforcement of farms characterized as those of cluster MCDMb-2.

5. Discussion and conclusions

The conclusions derived from this paper can be divided into methodological and empirical ones. Within the methodological conclusions, first, it is worth mentioning the *utility of the set of techniques available to construct global sustainability indicators*. Those methods allow the aggregation of a multidimensional set of indicators into a unique composite indicator, which can facilitate the understanding of complex concepts as the agricultural sustainability. Thus, composite indicators could be considered as key elements to guide the public decision-making. Furthermore, the use of these indices have other advantages as a tool to: a) improve public communication (mass media and the whole society), b) make comparisons between different geographical scopes (agricultural systems, nations, etc.), and c) make comparisons in different time scopes, in order to define tendencies.

In any case, the main disadvantages of these composite indicators should be also addressed. In this sense the two most important problems are associated with these indices are: a) the subjectivity associated to the selection of the aggregation method and b) the implementation of additive aggregation methods that allows compensability between the different dimensions of sustainability (commensurability).

Regarding the first issue pointed out above (*subjectivity in the selection of the aggregation method*), the empirical applications developed confirm that the results of the indices can vary depending on the aggregation method selected. Thus, it should be highlighted the relevance of the joint application of diverse techniques to a particular case study, in order to find consistent results before drawing final conclusions.

Within this point, the weighting process should be also considered. Usually it has been criticized the subjectivity in the weighting estimation, as a key element that can bias results. In this sense, the use of "objective" techniques has been suggested in order to avoid any particular assumption about the weights assigned to the different indicators. The method to construct indices on the basis of the PCA is a good example of this (weights are derived from the covariance matrix of indicators). However, if it is assumed that sustainability is a social construction process, it is compulsory to introduce the social preferences in the analysis. Only doing so the sustainability index takes into account socially acceptable trade-offs between economic, social and environmental indicators. Thus, without underestimating the usefulness of the aggregation method based on the PCA, the calculation of composite indicators using AHP or MCDM technique can be considered more suitable for the policy analysis of the agricultural sustainability.

A comparison of the empirical results obtained evidences that results depend on the *degree of commensurability* assumed by each method. The larger differences have been found between those methods that allow total commensurability (additive methods, such as PCA and AHP) and the GSI-MCDM for λ=0, the one where non-commensurability between indicators is assumed. Regarding this point, some authors (Hansen, 1996; Bockstaller et al., 1997; Morse et al., 2001; Ebert and Welsch, 2004; Munda, 2005) have criticized the additive aggregation methods because they consider that trade-offs between attributes (commensurability) are incompatible with the concept of sustainability. In this sense, it is reasonable to opt by a technique that allows partial commensurability, as the GSI-MCDM for 0<λ<1. In any case, the selection of the most suitable value of λ (degree of commensurability) is an opened issue to be discussed in further research works.

To sum up, it should be noted that although a joint analysis of different aggregation methods to construct composite indicators is advised, the most interesting one is the GSI-MCDM for 0<λ<1. This predilection is based on: a) the possibility of considering social preferences and b) the assumption of partial commensurability between indicators. Within this aggregation technique, the degree of commensurability (value of λ) should be selected by users (i.e., policy decision-makers).

Regarding the results obtained in this research it is also interesting to remark the empirical conclusions derived for the case study analysed. First, it can be noted that a *hierarchy of policy scenarios regarding their sustainability* have been established. In this sense, although the indices GSI-AHP and GSI-MCDM for λ>0 have not shown any significant differences, because of the results of GSI-PCA and the GSI-MCDM for λ=0 we can point out the Luxemburg agreement (LA) scenario as the most sustainable one. This conclusion supports the last reform of the CAP. Opposite, the least sustainable agricultural policy scenario is world markets (WM), where a further liberalization of international markets and the withdrawal of public subsidies for the agricultural sector are assumed.

Second, the *heterogeneity of farms* in the case study regarding their sustainability is also worth to be mentioned. In this line, it can be noted that similar results have been obtained for all aggregation techniques employed. These results show that full-time farmers, with small-medium size holdings and sowing higher value-added crops (maize, sugar-beet and vegetables) are the most sustainable ones for all policy scenarios.

These empirical results confirm the usefulness of sustainability composite indicators to guide policy decision-making in the agricultural sector. For the irrigated agriculture in the Duero basin it can be concluded that any public policy willing to promote sustainability should be based on: a) a market policy that supports the farmers' incomes by using partial decoupling subsidies, and b) a structural policy that promotes small and medium multifunctional farms.

Acknowledgements

This research was supported by the Spanish Ministry of Education and Science (MEC) through the project FUTURPAC (AGL2006-05587-C04-01).

References

Aguarón, J. and Moreno-Jiménez, J.M. (2000). Stability intervals in the analytic hierarchy process. *European Journal of Operational Research* 125:114-133.

Andreoli, M. and Tellarini, V. (2000). Farm sustainability evaluation: methodology and practice. *Agriculture, Ecosystems and Environment* 77(1):43-52.

Bazzani, G.M., Viaggi, D., Berbel J., López, M.J. and Gutiérrez, C. (2004). A methodology for the analysis of irrigated farming in Europe. In J. Berbel and C. Gutiérrez (eds.), *Sustainability of European Agriculture under Water Framework Directive and Agenda 2000*. European Commission. Brussels, pp. 49-66.

Berbel, J. and Gutiérrez, C. (eds.) (2004). *Sustainability of European Agriculture under Water Framework Directive and Agenda 2000*. European Commission. Brussels.

Bockstaller, C., Girardin P. and van der Werf, H.M.G. (1997). Use of agro-ecological indicators for the evaluation of farming systems. *European Journal of Agronomy* 7(1-3):261-270.

Diaz-Balteiro, L. and Romero, C. (2004). In search of a natural systems sustainability index. *Ecological Economics* 49:401-405.

Easley, R., Valacich, J. and Venkataramanan, M. (2000). Capturing group preferences in multicriteria decision. *European Journal of Operational Research* 125:73-83.

Ebert, U. and Welsch, H. (2004). Meaningful environmental indices: a social choice approach. *Journal of Environmental Economics and Management* 47(2):270–283.

Forman, E. and Peniwati, K. (1998). Aggregating individual judgments and priorities with the Analytic Hierarchy Process. *European Journal of Operational Research* 108:165-169.

Freudenberg, M. (ed.) (2003). *Composite indicators of country performance: a critical assessment*. OECD, Paris.

Hediger, W. (1999). Reconciling 'weak' and 'strong' sustainability. *International Journal of Social Economics* 26 (7/8/9):1120-1143.

Gómez-Limón, J.A. and Atance, I. (2004). Identification of public objectives related to agricultural sector support. *Journal of Policy Modelling* 27(8-9):1045-1071.

Gómez-Limón, J.A. and Riesgo, L. (2004). Irrigation water pricing: Differential impacts on irrigated farms. *Agricultural Economics* 31(1):47-66.

Hansen, J.W. (1996). Is agricultural sustainability a useful concept?. *Agricultural Systems* 50(1):117-143.

Morris, J., Gómez, M., Vasileiou, K. and Berbel, J. (2004). WADI Scenario definition. In J. Berbel and C. Gutiérrez (eds.) *Sustainability of European Agriculture under Water Framework Directive and Agenda 2000*. European Commission. Brussels, pp. 39-47.

Morse, S., Mcnamara, N., Acholo, M. and Okwoli, B. (2001). Sustainability indicators: the problem of integration. *Sustainable Development* 9(1):1-15.

Munda, G. (2005). Measuring sustainability: a multi-criterion framework. *Environment, Development and Sustainability* 7(1):117-134.

Nardo, M, Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A. and Giovannini, E. (eds.) (2005a). *Handbook on constructing composite indicators: methodology and user guide*. OECD Statistics Working Paper. OECD, Paris.

Nardo, M, Saisana, M., Saltelli, A. and Tarantola, S. (eds.) (2005b). *Tools for composite indicators building*. Joint Research Centre-European Commission, Ispra (Italy).

OECD (eds.) (2001). *Environmental indicators for agriculture. Volume 3 - Methods and Results*. OECD, París.

Pirazzoli, C. and Castellini, A. (2000). Application of a model for evaluating the environmental sustainability of cultures in hill and mountain areas. The case of berries and fruit chestnut in Northern Italy. Agricultural *Economics Review* 1(1):57-70.

Riesgo, L. and Gómez-Limón, J.A. (2005). Análisis de escenarios de políticas para la gestión pública de la agricultura de regadío. *Economía Agraria y Recursos Naturales* 5(9): 81-114.

Riesgo, L. and Gómez-Limón, J.A. (2006). Multi-criteria policy scenarios analysis for public regulation of irrigated agriculture. *Agricultural Systems* 91(1-2):1-28.

Rigby, D., Woodhouse, P., Young, T. and Burton, M. (2001). Constructing a farm level indicator of sustainable agricultural practice. *Ecological Economics* 39(3):463-478.

Saaty, T.L. (eds.) (1980). *The Analytic Hierarchy Process*. McGraw, New York.

Sands, G.R. and Podmore, T.H. (2000). A generalized environmental sustainability index for agricultural systems. *Agriculture, Ecosystems & Environment* 79(1):29-41.

Stockle, C.O., Papendick, R.I., Saxton, K.E., Campbell, G.S. and van Evert, F.K. (1994). A framework for evaluating the sustainability of agricultural production systems. *American Journal of Alternative Agriculture* 9(1-2):45-50.

Stoorvogel, J.J., Antle, J.M., Crissman, C.C. and Bowen W. (2004). The tradeoff analysis model: integrated bio-physical and economic modeling of agricultural production systems. *Agricultural Systems* 80(1):43-66.

van Calker, K.J., Berentsen, P.B.M., Romero, C., Giesen, G.W.J. and Huirne, R.B.M. (2006). Development and application of a multi-attribute sustainability function for Dutch dairy farming systems. *Ecological Economics* 57(3):640-658.

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 and Environment* 93(1):131-145.

Yunlong, C. and Smit, B. (1994). Sustainability in agriculture: a general review. *Agriculture, Ecosystems and Environment* 49(2):299-307.

Zahedi, F. (1987). A utility approach to the Analytic Hierarchy Process. *Mathematical Modelling* 9:387-395.

Annex

Table A1 - Results of GSI-PCA and GSI-AHP for each farm-type and policy scenario

Farm-type	**PCA Method (*GSI-PCA*)**						**AHP Method (*GSI-AHP*)**					
	SQ	**LA**	**WM**	**GS**	**PE**	**LS**	**SQ**	**LA**	**WM**	**GS**	**PE**	**LS**
1	0.523	0.501	0.420	0.430	0.514	0.381	0.508	0.404	0.413	0.391	0.521	0.340
2	0.509	0.508	0.479	0.421	0.510	0.471	0.505	0.460	0.459	0.374	0.499	0.385
3	0.511	0.518	0.347	0.386	0.427	0.406	0.504	0.362	0.267	0.378	0.405	0.396
4	0.474	0.544	0.392	0.412	0.388	0.496	0.420	0.494	0.411	0.363	0.309	0.403
5	0.535	0.642	0.505	0.564	0.506	0.526	0.599	0.660	0.571	0.646	0.537	0.579
6	0.487	0.604	0.510	0.418	0.436	0.532	0.390	0.533	0.493	0.357	0.339	0.499
7	0.604	0.618	0.499	0.524	0.472	0.534	0.566	0.524	0.404	0.517	0.449	0.505
8	0.596	0.647	0.525	0.547	0.547	0.535	0.638	0.616	0.580	0.594	0.557	0.565
9	0.575	0.617	0.485	0.541	0.438	0.536	0.614	0.625	0.545	0.595	0.406	0.555
10	0.607	0.663	0.585	0.524	0.516	0.591	0.607	0.621	0.603	0.552	0.458	0.558
11	0.467	0.472	0.320	0.423	0.402	0.447	0.371	0.359	0.330	0.350	0.297	0.401
12	0.364	0.425	0.347	0.404	0.433	0.454	0.353	0.373	0.368	0.370	0.466	0.389
13	0.502	0.538	0.370	0.437	0.496	0.412	0.509	0.524	0.332	0.368	0.516	0.358
14	0.371	0.450	0.485	0.450	0.500	0.466	0.347	0.389	0.533	0.460	0.523	0.442
15	0.534	0.607	0.446	0.494	0.527	0.480	0.621	0.494	0.549	0.568	0.612	0.508
16	0.513	0.564	0.448	0.480	0.482	0.478	0.621	0.632	0.554	0.558	0.540	0.514
17	0.605	0.586	0.494	0.545	0.534	0.524	0.623	0.460	0.540	0.571	0.546	0.511
18	0.494	0.360	0.458	0.526	0.449	0.471	0.546	0.472	0.650	0.602	0.587	0.632
19	0.553	0.583	0.500	0.528	0.521	0.497	0.566	0.540	0.515	0.553	0.539	0.536
20	0.396	0.644	0.379	0.491	0.523	0.484	0.410	0.585	0.368	0.580	0.585	0.539
21	0.416	0.352	0.495	0.515	0.519	0.458	0.418	0.347	0.509	0.540	0.516	0.454
22	0.536	0.533	0.408	0.428	0.476	0.450	0.502	0.446	0.400	0.397	0.426	0.403
Total	**0.498**	**0.568**	**0.496**	**0.515**	**0.521**	**0.532**	**0.524**	**0.548**	**0.542**	**0.541**	**0.510**	**0.528**

Table A2 - Results of GSI-MCDM for each farm-type and policy scenario

Farm-type	**λ=0**						**λ=0.25**						**λ=0.5**						**λ=0.75**					
	SQ	**LA**	**WM**	**GS**	**PE**	**LS**	**SQ**	**LA**	**WM**	**GS**	**PE**	**LS**	**SQ**	**LA**	**WM**	**GS**	**PE**	**LS**	**SQ**	**LA**	**WM**	**GS**	**PE**	**LS**
1	0.000	0.003	0.000	0.004	0.008	0.000	0.127	0.103	0.103	0.101	0.136	0.085	0.254	0.203	0.207	0.197	0.265	0.171	0.381	0.304	0.310	0.294	0.393	0.256
2	0.004	0.007	0.000	0.004	0.007	0.000	0.129	0.120	0.115	0.096	0.130	0.096	0.254	0.234	0.230	0.189	0.253	0.192	0.380	0.347	0.344	0.281	0.376	0.289
3	0.003	0.000	0.000	0.005	0.008	0.007	0.128	0.090	0.067	0.098	0.108	0.105	0.253	0.181	0.134	0.192	0.207	0.202	0.379	0.271	0.200	0.285	0.306	0.299
4	0.000	0.004	0.000	0.000	0.000	0.003	0.105	0.126	0.103	0.091	0.077	0.103	0.210	0.249	0.206	0.182	0.154	0.203	0.315	0.372	0.308	0.272	0.232	0.303
5	0.000	0.012	0.000	0.009	0.002	0.009	0.150	0.174	0.143	0.169	0.136	0.151	0.300	0.336	0.286	0.328	0.270	0.294	0.449	0.498	0.429	0.487	0.404	0.436
6	0.000	0.000	0.000	0.002	0.005	0.000	0.098	0.133	0.123	0.091	0.088	0.125	0.195	0.267	0.246	0.180	0.172	0.250	0.293	0.400	0.369	0.268	0.255	0.375
7	0.006	0.009	0.000	0.000	0.003	0.003	0.146	0.138	0.101	0.129	0.115	0.129	0.286	0.267	0.202	0.259	0.226	0.254	0.426	0.396	0.303	0.388	0.338	0.380
8	0.013	0.012	0.000	0.006	0.010	0.008	0.169	0.163	0.145	0.153	0.147	0.147	0.325	0.314	0.290	0.300	0.284	0.286	0.482	0.465	0.435	0.447	0.421	0.425
9	0.012	0.006	0.000	0.010	0.009	0.008	0.162	0.161	0.136	0.156	0.108	0.144	0.313	0.315	0.273	0.302	0.208	0.281	0.463	0.470	0.409	0.448	0.307	0.418
10	0.009	0.008	0.000	0.000	0.007	0.007	0.158	0.161	0.151	0.138	0.120	0.145	0.308	0.314	0.301	0.276	0.232	0.283	0.457	0.468	0.452	0.414	0.345	0.420
11	0.000	0.000	0.000	0.000	0.000	0.000	0.093	0.090	0.083	0.087	0.074	0.100	0.186	0.179	0.165	0.175	0.149	0.200	0.279	0.269	0.248	0.262	0.223	0.301
12	0.000	0.000	0.000	0.000	0.006	0.000	0.088	0.093	0.092	0.093	0.121	0.097	0.177	0.187	0.184	0.185	0.236	0.194	0.265	0.280	0.276	0.278	0.351	0.292
13	0.003	0.003	0.000	0.000	0.007	0.001	0.130	0.133	0.083	0.092	0.134	0.090	0.256	0.263	0.166	0.184	0.262	0.179	0.382	0.394	0.249	0.276	0.389	0.268
14	0.000	0.000	0.000	0.000	0.009	0.003	0.087	0.097	0.133	0.115	0.138	0.113	0.174	0.195	0.267	0.230	0.266	0.223	0.260	0.292	0.400	0.345	0.394	0.332
15	0.000	0.003	0.000	0.000	0.000	0.000	0.155	0.126	0.137	0.142	0.153	0.127	0.311	0.249	0.275	0.284	0.306	0.254	0.466	0.372	0.412	0.426	0.459	0.381
16	0.000	0.000	0.000	0.000	0.000	0.000	0.155	0.158	0.138	0.140	0.135	0.129	0.311	0.316	0.277	0.279	0.270	0.257	0.466	0.474	0.415	0.419	0.405	0.386
17	0.010	0.006	0.000	0.002	0.008	0.009	0.163	0.120	0.135	0.144	0.142	0.134	0.317	0.233	0.270	0.286	0.277	0.260	0.470	0.346	0.405	0.429	0.412	0.385
18	0.000	0.000	0.000	0.000	0.000	0.000	0.137	0.118	0.162	0.151	0.147	0.158	0.273	0.236	0.325	0.301	0.294	0.316	0.410	0.354	0.487	0.451	0.440	0.474
19	0.000	0.003	0.000	0.000	0.005	0.004	0.141	0.138	0.129	0.138	0.139	0.137	0.283	0.272	0.257	0.277	0.272	0.270	0.424	0.406	0.386	0.415	0.406	0.403
20	0.001	0.004	0.000	0.003	0.007	0.001	0.104	0.149	0.092	0.147	0.151	0.136	0.206	0.295	0.184	0.292	0.296	0.270	0.308	0.440	0.276	0.436	0.440	0.404
21	0.000	0.000	0.000	0.000	0.005	0.008	0.104	0.087	0.127	0.135	0.133	0.120	0.209	0.173	0.254	0.271	0.260	0.231	0.313	0.260	0.381	0.406	0.388	0.342
22	0.004	0.007	0.000	0.004	0.005	0.002	0.128	0.117	0.100	0.102	0.111	0.102	0.253	0.226	0.200	0.200	0.216	0.202	0.378	0.336	0.300	0.299	0.321	0.303
Total	**0.007**	**0.009**	**0.000**	**0.007**	**0.008**	**0.010**	**0.136**	**0.144**	**0.136**	**0.140**	**0.134**	**0.140**	**0.266**	**0.278**	**0.271**	**0.274**	**0.259**	**0.269**	**0.395**	**0.413**	**0.406**	**0.407**	**0.385**	**0.399**

NOTE: Values of GSI-MCDM for λ=1 are not shown in this table because they are the same as those obtained for the GSI-AHP (see Table A1)