Data Analytics in Organic Farming: Impact on Environmental Sustainability

Cristhian Paz[1], Josune Sáenz[1] and Ana Ortiz-de-Guinea[2]
[1]Deusto Business School, University of Deusto, San Sebastián, Spain
[2] HEC Montréal, Montreal, Canada
cristhianpaz@deusto.es
josune.saenz@deusto.es
ana.ortiz-de-guinea@hec.ca

Abstract: The production of healthy food while preserving the environment constitutes one of the main challenges of the 21st century. Along these lines, organic farming has emerged as a farm management and food production system that encourages environmental sustainability. To enhance such sustainability, data analytics both as an asset and as a capability could play a substantial role. Indeed, data analytics could be used to interpret the past and predict the future and to make more timely or accurate decisions regarding the use and protection of natural resources. Using survey data from 119 Spanish organic farms whose digitization degree as reported by the farmer is above 0, and structural equation modeling based on partial least squares to test research hypotheses, we found that even though data analytics in organic farming is clearly underdeveloped, it still contributes to enhancing farms' environmental sustainability. Thus, investments in environmental data analytics appear to pay off.

Keywords: Documented knowledge, Data analytics, Environmental sustainability, Agriculture, Organic farming

1. Introduction

According to the European Commission (2022a), food systems cannot be resilient to crises such as the COVID-19 pandemic if they are not sustainable. Today, such systems account for nearly one-third of global greenhouse gas emissions, consume large amounts of natural resources, result in biodiversity loss, and provoke negative health impacts (European Commission 2022a). Therefore, the production of healthy food while preserving the environment constitutes one of the main challenges of the 21st century. Along these lines, organic farming has emerged as a farm management and food production system that encourages responsible usage of energy and natural resources; maintenance of biodiversity; preservation of regional ecological balances; enhancement of soil fertility; maintenance of water quality; and high standards of animal welfare (European Commission 2022b).

Data analytics (i.e., analyzing raw data to identify trends and draw conclusions that help decision making) could play a substantial role in ensuring organic farms' environmental sustainability. In 2019, the EU Member States signed a cooperation declaration entitled "A Smart and Sustainable Digital Future for European Agriculture and Rural Areas", which states that rural digitalization has the potential to increase efficiency on the farm, improve production, and contribute to making agri-food systems more sustainable from an economic, social, and environmental perspective. Regarding the latter aspect, a systematic literature review by Sharma et al (2020) found that machine learning—which usually enables data analytics functionality—could help manage soil properties; make irrigation decisions; detect pest infestations, diseases, and weeds before actual outbreaks; and manage site specific nutrients, among other applications.

Analytical tools, however, could range over very different degrees of sophistication. For instance, Carlos et al (2021), reported a situation in which, at the outset, a family-owned farm's foreman recorded crop irrigation attributes (such as field, crop, meter readings, on and off dates, etc.) with paper and pencil from the fields and then stored the data in physical binders. Such data was then transferred at the end of each week and each season from the written notes to a Microsoft Excel sheet for further analysis and archival. A step forward from this rather rudimentary and not real-time procedure involved implementing a water tracking system that collected field-specific irrigation data via a Google form on the foreman's mobile, then synchronized with other relevant crop data specific to the farm in a spreadsheet, synchronized with external weather data from the regional irrigation management and information system, and then updated the visualizations in the Excel workbook to help decision making.

Today, machines equipped with sensors and cameras offer further possibilities by capturing minute field-level data like soil moisture, leaf greenness, temperature, and yield, among others (Pham & Stack 2020) that could be analyzed with specific tools to obtain natural resource management insights. These new technologies pave the way to move farming's organizational culture away from experience-based management towards data-based management (Janc et al 2019), or at least towards a more balanced type of management.

Although different papers exist that describe existing technologies and recent developments that enable data collection and analysis, as well as their potential benefits in the agricultural domain (e.g., Bacco et al 2019; Moysiades et al 2021), research that assesses the degree of implementation of data analytics in organic farming for the management of natural resources (i.e., soil, water, plants, and energy) and its impact on environmental sustainability is missing. This constitutes an important research gap since farmers and government institutions need evidence on the effectiveness of these tools to make investment decisions.

This paper tries to overcome this gap by analyzing the degree of implementation of environmental data analytics both as an asset (i.e., up-to-date and easily accessible data records regarding key environmental aspects for farm management, whether digitized or not) and as a capability (in this case, the capability to align or realign—i.e., manage—natural resources based on the analysis of available and digitized data by means of software applications) and their impact on environmental sustainability in organic farming.

2. Theoretical background and hypothesis development

2.1 Theoretical background

Data analytics constitutes a knowledge resource and as such is part of the intellectual capital of the firm (Subramaniam & Youndt 2005; Peñalba-Aguirrezabalaga et al 2020). It involves both an asset and a capability dimension. As an asset, data analytics refers to collected data stored in documents and/or databases. The latter adds to the firm's structural capital as data remains in the company even though employees may leave (Bontis 1998). However, data analytics as a capability refers to the ability to analyze collected data using software to generate relevant knowledge for running the business (Grover et al 2018). It connects both with structural and human capital, as it depends both on the available software (i.e., structural capital) and on the individual's ability to use the software, interpret the results obtained, and make decisions (i.e., human capital).

This paper focuses on environmental data analytics or data analytics for the specific purpose of natural resource management in organic farming, with data availability and data digitization representing the asset dimension, and data analytic enabled environmental nature resource management representing the capability one. More precisely, data availability is the extent to which there are up-to-date and easily accessible records regarding the quality, use, and protection of (i.e., treatments applied to) natural resources, while digitization deals with the fact that such data could be digitized or not, as described in the example shown in the introduction section. On the other hand, environmental data analytics as a capability refers to the extent to which analytical software applications are used by organic farmers to generate diagnosis and recommendations regarding the use and protection of the farm's natural resources. As such, it constitutes a dynamic capability as it facilitates decision-making regarding changes in the natural resource base of the farm (Helfat et al 2007) based on insights and recommendations generated by the data analytic software.

To theorize about the relations of the elements of data analytics among themselves (i.e., data availability, data digitization, and data analytic enabled environmental data analytic capability), as well as with environmental sustainability, we take an information value chain theory perspective. This perspective provides mechanisms to tease apart the value added by each data analytic element in the information chain (Coiera 2019). Thus, this paper argues that data turns into environmental value (i.e., environmental sustainability) by following a set of steps, each adding value, that include data collection and storage, data digitization, and data analytic enabled knowledge generation for decision making regarding the organic farm's natural resources. In this context, the 'end' value of the chain is environmental sustainability, which according to the natural-resource-based view of the firm (Hart 1995; Hart & Dowell 2011), involves (a) pollution prevention (i.e., reducing waste and residues to a minimum); (b) product stewardship (i.e., lower product life cycle cost by minimizing natural resource consumption); and (c) sustainable development (i.e., producing in a way that can be maintained indefinitely into the future by protecting natural resources). This is summarized in Figure 1 in the form of a research model of which hypotheses are explained next.

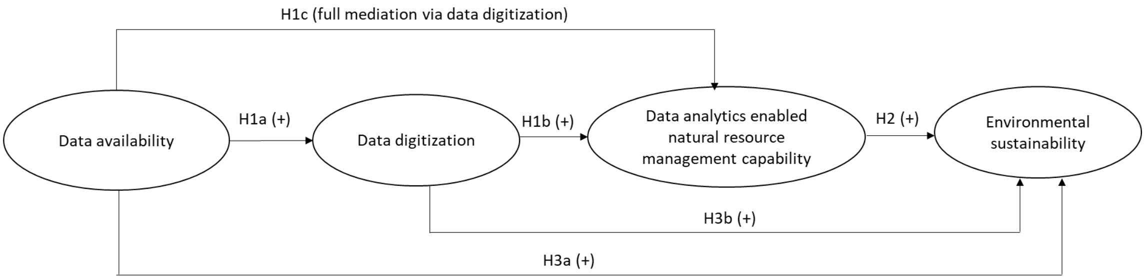

Figure 1: Research model

2.2 Hypothesis development

The capability to make decisions on natural resource usage and protection through the application of analytical software (i.e., data analytics enabled natural resource management capability) cannot be deployed without data. In other words, the insights generated by the application of data analytic tools is dependent upon the available data (Popovič & Habjan 2012). As pointed out in the introduction section, data may be collected manually and then stored on a physical device (e.g., a folder or binder) before being uploaded into a computer system for analysis or may be directly uploaded into an information system via sensors, thus digitally stored from the very outset. Accordingly, the following hypotheses can be formulated:

H1a: Data availability is positively related to data digitization.
H1b: Data digitization is positively related to data analytics enabled natural resource management capability.
H1c: The relationship between data availability and data analytics enabled natural resource management capability is fully mediated by data digitization (in other words, the direct relationship between data availability and data analytics enabled natural resource management capability is non-significant, while the indirect relationship is positive and significant).

Insights gained through data analytics regarding the quality, use, and protection of (i.e., treatments applied to) natural resources (i.e., soil/plants, water, and energy) would result in decisions that contribute to the enhancement of farms' environmental sustainability. For instance, soil and nutrient management software should lead to fertilization decisions that help preserve soil quality parameters while being in line with good environmental care; applications for identifying and controlling crops' pests and diseases should facilitate treatment decisions that help keep them under control; irrigation software should facilitate reducing water consumption to the strict minimum; likewise, energy management software exists that tracks peak energy usage and demand charges down to the meter and thus facilitates energy savings; finally, waste management applications are available that help reduce the amount of waste produced of different types and foster such waste to be reused. Based on the above, the following hypothesis is formulated:

H2: Data analytics enabled natural resource management capability is positively related to environmental sustainability.

Finally, the last two hypotheses acknowledge that documented and stored data (whether digitized or not) can also contribute to environmental sustainability without the need for digitization nor data analytic enabled resource management capability. The farmer can still be making good decisions regarding the farm's natural resources based on his/her assessment of non-digitized data, thus contributing to the environmental sustainability of his/her farm. Likewise, digitization is likely to also facilitate environmental sustainability without the need for the data analytic enabled capability. Therefore, the following hypotheses are suggested:

H3a: Data availability is positively related to environmental sustainability.
H3b: Data digitization is positively related to environmental sustainability.

3. Research methods

We analyze the degree of implementation of environmental data analytics and its influence on environmental sustainability using survey research data from a set of family-owned organic farms in Spain. Spain is the second European country with the largest area dedicated to organic agriculture behind France (2.5 million hectares vs. 2.4 million, respectively) (IFOAM 2022). The target population of the research was identified by means of the

REGOE (*Registro General de Operadores Ecológicos de España*), which is the Spanish General Register of Organic Operators. In all, 28,883 vegetable producers were identified whose farms were family-owned. From this population, a sample of 406 producers was obtained that respected regional and gender proportions. The choice was then made to focus on those producers whose main production was crop-based (358 producers) and that had all the elements to be analyzed in the research model (119 producers). Here, the main constrain referred to data digitization. As explained, only 119 producers were found that had digitized data (i.e., the data digitization degree as reported by the farmer was above 0 only in 119 cases), which represents just one third of the producers whose main production is crop-based.

A questionnaire was designed to gather information about the variables included in the research and was administered by phone by a specialized company. Total anonymity was guaranteed, and personalized reports of results were offered to encourage participation. All questionnaires were answered by people who actively participated in the management of the farm. Of these, 96 people (80.7%) were owners of the farm, and the remaining 23 (19.3%) while not being owners, were actively involved in decision making, planning and organization of day-to-day activities. Each farm had 1.78 permanent employees on average.

Measures corresponding to research variables (i.e., data availability, data digitization, data analytics enabled natural resource management capability, environmental sustainability, and control variables—size of the farm or agricultural production unit, and age of agroecological practices) are shown in Table 1. The data availability construct is formed by 5 indicators referring to data inputs regarding quality parameters of natural resources, natural resource usage, and resource protection (i.e., treatments applied). Consistent with this, data digitization measures the extent to which the above data is digitalized. On the other hand, data analytics enabled natural resource management capability is measured by 7 items capturing the extent to which the software allows the farmers to generate several types of diagnosis and decision recommendations regarding the natural resources of the farm. Finally, environmental sustainability encompasses 7 indicators that correspond to the 3 dimensions considered within the context of the natural-resource-based view (Hart 1995; Hart & Dowell 2011), as already explained in the theoretical background section.

In this case, indicators build up (or define) the conceptual variables. In other words, each research variable is conceived as a combination of different elements that make it up and that may not necessarily correlate. When this is the case, a composite measurement model applies (Henseler 2017). In composite measurement, latent variable scores are linear combinations (i.e., weighted composites) of the indicators making up the variables without error term. Two possibilities exist to estimate indicators' weights: mode "A" (i.e., correlation-based) and mode "B" (i.e., ordinary least squares regression-based) (Rigdon, 2016). According to Sarstedt et al (2016), mode "A" estimation performs better when sample sizes are small and R^2 values are small to medium, as is the case in this research for some of the dependent variables (see Table 2). Thus, mode "A" estimation was used.

Giving the emerging/formative nature of research variables, research hypotheses were analyzed using Structural Equation Modelling based on Partial Least Squares (SmartPLS 3.3.5 software) (Ringle, Wende, & Becker 2015).

4. Research findings

4.1 Descriptive analysis

Prior to evaluating the measurement and the structural model, a descriptive analysis was carried out to portray the degree of development of data analytics in the organic farms under study. As can be observed in Table 1, and as far as data analytics as an asset is concerned, on a scale from 0 to 10, digitization of environmental data tends to be 1 point below its documentation degree (i.e., data availability in any form), except for soil quality parameters and treatments applied, where the difference between documented data and digitized data is almost 2 points. Moreover, it should be noticed that data regarding treatments applied constitutes the only item in which the degree of documentation and digitization is well above 5. Therefore, it can be concluded that environmental data documentation and digitization tend to be quite low. The same happens with data analytics enabled natural resource management capability: on a scale from 0 to 10, average scores for each item are well below 5, although it should be noticed that standard deviations tend to be quite large, both in the case of data analytics as an asset and in the case of data analytics as a capability. Environmental sustainability items, however, are clearly above 5 in all cases.

4.2 Measurement model evaluation

In composite measurement, traditional assessment methods aimed at proving that the indicators chosen to operationalize the concepts derive from the same underlying phenomenon are not applicable. In this case, concepts emerge as the combination of the different elements included in the definition and thus measurement model assessment should guarantee that the indicators capture the essence of the conceptual variables (i.e., "conceptual fidelity").

As can be observed in Table 1, indicators chosen for both data availability and data digitization (i.e., environmental data analytics as an asset) include quality, usage, and protection parameters regarding main natural resources used in the farm (i.e., soil/plants, water, and energy), while the indicators chosen for data analytics enabled natural resource management capability include different diagnosis aspects and decision recommendations regarding the same natural resources which are facilitated by agricultural analytical tools available in the market. Thus, in the case of these variables, conceptual fidelity is quite straightforward. Finally, as far as environmental sustainability is concerned, as already explained, indicators included correspond to the three dimensions suggested by the natural-resource-based view: (a) pollution prevention (i.e., reducing waste and residues to a minimum: ENVSUST7); (b) product stewardship (i.e., lower product life cycle cost by minimizing natural resource consumption: ENVSUST1 and ENVSUST2); and (c) sustainable development (i.e., producing in a way that can be maintained indefinitely into the future by protecting natural resources: ENVSUST3, ENVSUST4, and ENVSUST5).

Collinearity between indicators (measured in terms of VIF values) is another aspect to be analyzed. However, this is only relevant in mode "B" (i.e., regression-based) estimation, as in such case collinearity (i.e., VIF values being above 3) can affect indicator's weights and could cause reversed signs. When this is the case, the researcher should consider replacing mode "B" estimation with mode "A", which is not affected by collinearity (Rigdon, 2016). In our case, such decision has been made beforehand, due to the rather small level of variance explained (i.e., R^2) of some of the dependent variables (see Table 2) and the relatively small sample size (119 producers). Such choice involves that indicators' weights are calculated in terms of their absolute relevance instead of in terms of their relative relevance vis-à-vis other indicators in the same construct when it comes to making up the latent variable score to maximize the amount of variance explained of the dependent variables.

Table 1: Measurement model evaluation

Constructs and measures	*Item wording*	*Mean (or %)*	*STDEV*	*VIF*	*Weight*	*p-val.*
Data availability (ENVDAV) Mode "A" composite	Please rate from 0 to 10 (0 = Not at all; 10 = Very much) the degree to which the agricultural production unit of reference has up-to-date and easily accessible records on the aspects listed below:					
ENVDAV1	The evolution of soil quality parameters.	4.983	4.079	1.764	0.286	0.000
ENVDAV2	The evolution of water quality parameters.	3.958	4.204	1.912	0.339	0.000
ENVDAV3	Irrigation statistics: irrigation days, irrigation time, water consumption...	4.697	4.378	1.576	0.289	0.000
ENVDAV4	Energy consumption statistics.	3.832	4.163	1.524	0.266	0.000
ENVDAV5	Treatments applied: dates, actions carried out, doses applied...	8.456	2.887	1.159	0.181	0.000
Data digitization (ENVDDIG) Mode "A" composite	Please rate from 0 to 10 (0 = Not at all; 10 = Completely) the degree to which the information records previously assessed are digitized:					
ENVDDIG1	The evolution of soil quality parameters.	3.269	4.293	1.960	0.262	0.000
ENVDDIG2	The evolution of water quality parameters.	2.840	4.019	1.996	0.292	0.000
ENVDDIG3	Irrigation statistics: irrigation days, irrigation time, water consumption...	3.353	4.269	1.529	0.308	0.000
ENVDDIG4	Energy consumption statistics.	2.882	4.061	1.526	0.300	0.000
ENVDDIG5	Treatments applied: dates, actions carried out, doses applied...	6.538	4.107	1.276	0.186	0.000
Data analytics enabled natural resource	Please rate from 0 to 10 (0 = Not at all; 10 = Very much) the degree to which the computer applications and tools available in the agricultural production unit allow you to:					

Constructs and measures	*Item wording*	*Mean (or %)*	*STDEV*	*VIF*	*Weight*	*p-val.*
management capability (ENVRAC) Mode “A” composite						
ENVRAC1	Identify nutritional deficiencies in crops (i.e., lack of substances necessary for proper plant development).	3.042	3.842	3.365	0.126	0.000
ENVRAC2	Assess the impact of pests and diseases.	3.218	3.887	4.492	0.154	0.000
ENVRAC3	Make care and treatment decisions.	3.739	4.047	4.084	0.159	0.000
ENVRAC4	Perform calculations regarding energy efficiency and energy savings.	3.731	3.995	4.400	0.195	0.000
ENVRAC5	Evaluate irrigation practices and calculate water efficiency and water footprint.	3.445	4.045	5.636	0.201	0.000
ENVRAC6	Generate recommendations on water consumption and management.	3.126	3.888	5.476	0.214	0.000
ENVRAC7	Calculate carbon footprint.	2.000	3.259	2.124	0.109	0.000
Environmental sustainability (ENVSUST) Mode “A” composite	Please rate from 0 to 10 (0 = Strongly disagree; 10 = Strongly agree) your degree of agreement or disagreement with the following statements regarding the agricultural production unit of reference:					
ENVSUST1	Our energy consumption has reached an optimum level: there is no other technique or action that will allow us to reduce it further.	6.227	3.211	1.314	0.003	0.490
ENVSUST2	Our water consumption has reached an optimal level: there is no other technique on the market or type of action that will allow us to reduce it further.	7.042	3.281	1.480	0.197	0.017
ENVSUST3	Our fertilizer supply to the soil is always in line with the requirements of good environmental care.	9.286	1.540	1.263	0.163	0.108
ENVSUST4	We have a good soil quality: all parameters are at adequate levels.	8.025	2.392	1.831	0.334	0.000
ENVSUST5	Pests and/or diseases are adequately controlled.	7.403	2.522	1.638	0.299	0.000
ENVSUST6	The biodiversity of our agricultural production unit is very high.	7.798	2.693	1.682	0.309	0.000
ENVSUST7	The volume of waste and residues generated is minimal.	8.529	1.913	1.517	0.152	0.058
Control variables						
Size of the agricultural production unit (SIZE)	Natural logarithm of the number of hectares of the agricultural production unit.	74.84*	395.54	N/A	N/A	N/A
Age of agroecological practices (AGE)	Year of introduction of agroecological practices.	2,010	6.35	N/A	N/A	N/A

*Average number of hectares devoted to crops and pastures.

4.3 Structural model evaluation

To test the significance and strength of the research hypotheses, we used a one-tailed 5,000 subsample BCA bootstrap (Hair et al 2017). Table 2 shows the results obtained.

Table 2: Structural model evaluation

	Effects	*STDEV*	*t statistics*	*p-values*	*10%*	*90%*
Direct relationships regarding data digitization (R^2 = 63.8%)						
Size	**0.109**	**0.054**	**2.031**	**0.021**	**0.040**	**0.176**
Data availability	**0.788**	**0.041**	**19.302**	**0.000**	**0.727**	**0.834**
Direct relationships regarding data analytics enabled natural resource management capability (R^2 = 15.5%)						

	Effects	*STDEV*	*t statistics*	*p-values*	*10%*	*90%*
Size	-0.086	0.083	1.032	0.151	-0.195	0.020
Data availability	0.067	0.151	0.442	0.329	-0.126	0.261
Data digitization	**0.340**	**0.155**	**2.201**	**0.014**	**0.125**	**0.522**
Direct relationships regarding environmental sustainability (R^2 = 13.5%)						
Size	-0.057	0.118	0.479	0.316	-0.207	0.097
Age of agroecological practices	-0.032	0.119	0.267	0.395	-0.164	0.139
Data availability	**0.352**	**0.168**	**2.091**	**0.018**	**0.090**	**0.510**
Data digitization	-0.134	0.181	0.741	0.230	-0.347	0.094
Data analytics enabled natural resource management capability	**0.194**	**0.095**	**2.046**	**0.020**	**0.042**	**0.287**
Indirect relationship between data availability and data analytics enabled natural resource management capability						
Indirect relationship via data digitization	**0.268**	**0.124**	**2.166**	**0.015**	**0.101**	**0.414**
Total degree of association between data availability and data analytics enabled natural resource management capability (Direct + Indirect)	**0.335**	**0.088**	**3.804**	**0.000**	**0.209**	**0.435**
Indirect relationships between data availability and environmental sustainability						
Indirect relationship via data digitization (1)	-0.105	0.145	0.725	0.234	-0.279	0.072
Indirect relationship via data analytics enabled natural resource management capability (2)	0.013	0.034	0.385	0.350	-0.019	0.063
Indirect relationship via data digitization and via data analytics enabled natural resource management capability (3)	**0.052**	**0.038**	**1.367**	**0.086**	**0.012**	**0.107**
Total indirect relationship (1+2+3)	-0.040	0.140	0.289	0.386	-0.222	0.124
Total degree of association between data availability and environmental sustainability (Direct + Indirect)	**0.312**	**0.101**	**3.074**	**0.001**	**0.116**	**0.400**
Indirect relationship between data digitization and environmental sustainability						
Indirect relationship via data analytics enabled natural resource management capability	**0.066**	**0.048**	**1.381**	**0.084**	**0.015**	**0.135**
Total degree of association between data digitization and environmental sustainability (Direct + Indirect)	-0.068	0.179	0.337	0.353	-0.288	0.157

As can be observed in Table 2, data availability is positively and significantly related to data digitization (β = 0.788), while the latter is also significantly and positively related to data analytics enabled natural resource management capability (β = 0.340). However, there is no direct significant relationship between data availability and natural resource management capability facilitated by data analytics, meaning that environmental data availability facilitates natural resource management capability based on data analytics inasmuch it allows ulterior data digitization, as the positive and highly significant indirect relationship via environmental data digitization points out (β = 0.268). Hence, hypothesis H1a, H1b, and H1c are supported.

On the other hand, data analytics enabled natural resource management capability is positively and significantly related to environmental sustainability (β = 0.194). Therefore, hypothesis H2 is supported. Furthermore, data availability is directly and positively related to environmental sustainability (β = 0.352), meaning that hypothesis H3a is supported. Thus, natural resource management capability based on documented data and personal judgement appears to have a greater effect on environmental sustainability than data analytics enabled natural resource management capability. However, there is no direct relationship between data digitization and environmental sustainability, but only indirect. As a result, hypothesis H3b is not supported. Figure 2 provides a visual summary of the results.

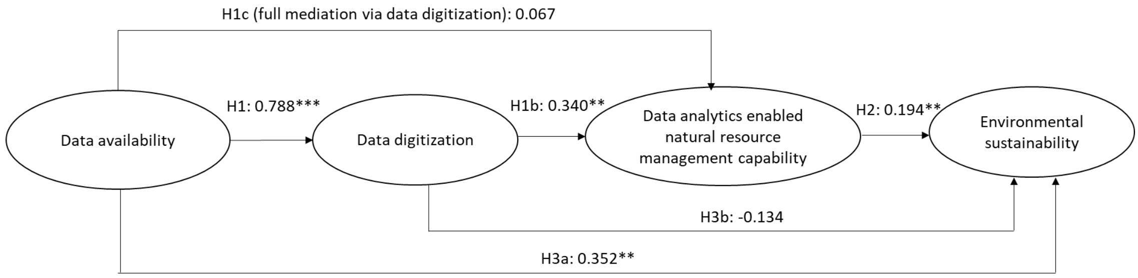

Figure 2: Results of the structural model

5. Discussion and conclusions

The research carried out analyzed the degree of implementation of environmental data analytics in family-owned organic farms in Spain and its influence on environmental sustainability. In this regard, the results indicate a positive relationship between data analytics and environmental sustainability, even though the mere fact of having environmental documented data that allows decision making based on personal judgement appears to be more influential than natural resource management capability based on ICT tools for data analytics. The prominent role of documented data that informs decision making and improves performance (in this case, environmental sustainability) is in line with the long tradition of studies in the intellectual capital domain that analyze the structural capital-performance link. However, it is not negligible that despite the low implementation of data analytics the relationship between data analytics enabled natural resource management capability and environment sustainability proved to be positive and significant, which makes think that this relationship could become stronger in the future, as the degree of implementation of data analytics increases.

Several reasons may explain the low levels of data analytic adoption by organic farms. While a clear one is the small size of the farms in the sample, others might be related to the demographics of the farmers (for instance, the average age of respondents was 51) and to infrastructure connectivity issues in rural areas. Moreover, the compatibility of digitalization with approaches such as organic farming is a matter of debate among farmers (Plumecocq et al 2018; Van Hulst et al 2020). One farmers' concern is the loss of technological sovereignty: while agriculture, peasant livestock, and artisanal fisheries have always produced and shared their own technologies, most new technologies are generated outside the system in which they are applied. This implies that those who use them must invest part of their income to pay for them. Furthermore, the types of commercial relationships established are often not punctual, but tie the producers down after the initial investment with ongoing costs of purchasing materials, software updates, maintenance services, etc. (Oteros-Rozas 2020).

Finally, this study has some limitations that warrant future research. For example, while our research method (i.e., survey) allowed us to find a low level of digitization among organic farms, qualitative methods such as case studies could help in explaining this finding. Furthermore, qualitative case studies could also be used to deepen into the specific tools used for data analytics by family-owned organic farms, the benefits they obtain from them, and the difficulties they face.

Acknowledgements

This paper is part of the NERSO project (Nuevos Modelos de Negocio para un Entorno Rural Sostenible / New Business Model for a Sustainable Rural Environment) which has been funded by the Science and Innovation Ministry of the Spanish Government (project reference: PID2020-114658RB-I00).

References

Bacco, M., Barsocchi, P., Ferro, E., Gotta, A. and Ruggeri, M. (2019) "The Digitisation of Agriculture: A Survey of Research Activities on Smart Farming", *Array*, Vol 3-4, pp 1-11.

Bontis, N. (1998) "Intellectual Capital: An Exploratory Study That Develops Measures and Models", *Management Decision*, Vol 36, No. 2, pp 63-76.

Carlos, A., Salinthone, X., Nguyen, M. and Jacobs, A. (2021) "Sustainable Agriculture Through Data Analytics", *The CSU Journal of Sustainability and Climate Change*, Vol 1, Article 2, pp 1-12.

Coiera, E. (2019) “Assessing Technology Success and Failure Using Information Value Chain Theory”, *Studies in Health Technology and Informatics*, Vol 263, pp 35-48.

European Commission (2019) *Declaration – A Smart and Sustainable Digital Future for European Agriculture and Rural Areas*, Brussels.

European Commission (2022a) “Farm to Fork Strategy for a Fair, Healthy and Environmentally-Friendly Food System)”, [online], https://ec.europa.eu/food/horizontal-topics/farm-fork-strategy_en.

European Commission (2022b) “Organics at a Glance”, [online], https://ec.europa.eu/info/food-farming-fisheries/farming/organic-farming/organics-glance_en.

Grover, V., Chiang, R. H., Liang, T. P. and Zhang, D. (2018) “Creating Strategic Business Value from Big Data Analytics: A Research Framework”, *Journal of Management Information Systems*, Vol 35, No. 2, pp 388-423.

Hair, J. F., Hult, G. T. M., Ringle, C. M. and Sarstedt, M. (2017) *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)* (2nd ed.), Sage Publications, Thousand Oaks.

Hart, S. L. (1995) “A Natural-Resource-Based View of The Firm”, *Academy of Management Review*, Vol 20, pp 986-1014.

Hart, S. L. and Dowell, G. (2011) “A Natural-Resource-Based View of the Firm: Fifteen Years after”, *Journal of Management*, Vol 37, No. 5, pp 1464-1479.

Helfat, E., Finkelstein, S., Mitchell, W., Peteraf, M. A., Singh, H., Teece, D. J. and Winter, S. G. (2007) *Dynamic Capabilities: Understanding Strategic Change in Organizations*, Blackwell Publishing, Oxford.

Henseler, J. (2017) “Bridging Design and Behavioral Research with Variance-Based Structural Equation Modeling”, *Journal of Advertising*, Vol 46, No. 1, pp 178-192.

IFOAM (2022) “New Organic Data: Growth Continued in 2020 but Policy Support on Statistics and CAP Essential to Reach EU targets”, [online], https://www.organicseurope.bio/news/new-organic-data-growth-continued-in-2020-but-policy-support-on-statistics-and-cap-essential-to-reach-eu-targets/.

Janc, K., Czapiewski, K. and Wójcik, M. (2019) “In the Starting Blocks for Smart Agriculture: The Internet as a Source of Knowledge in Transitional Agriculture”, *Wageningen Journal of Life Sciences*, Vol 90-91, pp 1-12.

Moysiadis, V., Sarigiannidis, P., Vitsas, V. and Khelifi, A. (2021) “Smart Farming in Europe”, *Computer Science Review*, Vol 39, pp 1-22.

Oteros-Rozas, E. (2020) “Digitalización y Revolución Agraria: Un Debate Complejo”, Soberanía Alimentaria, Biodiversidad y Culturas, Vol 39, https://soberaniaalimentaria.info/numeros-publicados/75-numero-39/805-digitalizacion-y-revolucion-agraria-un-debate-complejo.

Peñalba-Aguirrezabalaga, C., Sáenz, J. and Ritala, P. (2020) “Marketing-Specific Intellectual Capital: Conceptualization, Scale Development, and Empirical Illustration”, *Journal of Intellectual Capital*, Vol 21, No.6, pp 947-984.

Pham, X. and Stack., M. (2018) “How Data Analytics Is Transforming Agriculture”, *Business Horizons*, Vol 61, pp 125-133.

Popovič, A. and Habjan, A. (2012) “Exploring the Effects of Information Quality Change in Road Transport Operations”, *Industrial Management & Data Systems*, Vol 112, No. 8–9, pp 1307–1325.

Plumecocq, G., Debril, T., Duru, M., Magrini, M. B., Sarthou, J. P. and Therond, O. (2018) “The Plurality of Values in Sustainable Agriculture Models: Diverse Lock-in and Coevolution Patterns”, *Ecology and Society*, Vol 23, No. 1., https://doi.org/10.5751/ES-09881-230121.

Rigdon, E. E. (2016) “Choosing PLS Path Modeling as Analytical Method in European Management Research: A Realist Perspective”, *European Management Journal*, Vol 34, No. 6, pp 598-605.

Ringle, C. M., Wende, S. and Becker, J.-M. (2015) *SmartPLS 3*, Bönningstedt: SmartPLS GmbH.

Sarstedt, M., Hair, J. F., Ringle, C. M., Thiele, K. O. and Gudergan, S. P. (2016) “Estimation Issues with PLS and CBSEM: Where the Bias Lies!”, *Journal of Business Research*, Vol 69, No. 10, pp 3998-4010.

Sharma, R., Kamble, S. S., Gunasekaran, A. and Kumar, V. (2020) “A Systematic Literature Review on Machine Learning Applications for Sustainable Agriculture Supply Chain Performance”, *Computers and Operations Research*, Vol 119, pp 1-17.

Subramaniam, M. and Youndt, M. A. (2005), “The Influence of Intellectual Capital on the Types of Innovative Capabilities”, *The Academy of Management Journal*, Vol 48, No. 3, pp 450-463.

Van Hulst, F., Ellis, R., Prager, K. and Msika, J. (2020) “Using Co-Constructed Mental Models to Understand Stakeholder Perspectives on Agroecology”, *International Journal of Agricultural Sustainability*, Vol 18, No. 2, pp 172-195.