Organic farming in Italy: comparison and integration among sources for improving data consistency

Roberto Gismondi[1], Chiara Gnesi[1], Pietro Nurzia[1]

[1]Italian National Statistical Institute, Italy

Abstract

Organic farming is an important feature of agricultural holdings in the European Union. It indicates the propensity of farmers to guarantee sustainable agriculture. The European Commission has set the target of at least 25% of the EU's agricultural land under organic farming by 2030. In Italy, there are three main data sources on organic farming. The first one derives from administrative data managed by the Ministry of Agriculture, used for fulfil the requests of the ESS Agreement on Organic Production Statistics ESSC 2020/42/6/EN. The second source derives from the Integrated Farm Statistics Regulation IFS (EU) 2018/1091, which ruled the last agriculture census. The census 2020 collected data on organic farming directly from farms. The third source is the Farm Accountancy Data Network (FADN), established by Regulation (EC) 1217/2009. The FADN collects data on organic production: yield per hectare and quantity of milk per cow. EUROSTAT and DG AGRI request consistency among administrative data and those ruled by the IFS and the FADN Regulations. As regards 2020, through record linkage, the administrative microdata was compared with the census microdata, as regards the feature "the farm is organic or not". On average, discordant farms are larger and are mainly located in the Northeast. Furthermore, organic surfaces were multiplied by the correspondent yield coefficients supplied by FADN, to estimate organic crop production, as requested by the SAIO Regulation (EU) 2022/2379.

Keywords: Administrative data, Agriculture, Census, Consistency, Linkage

1. Introduction

Organic farming is an agricultural production system defined and regulated at the community level by EU regulations n 848/2018 and n.1165/2021. It does not use synthetic chemical products to fertilize land or to fight weeds, animal parasites, and plant diseases. The elimination of the use of chemical fertilizers, insecticides herbicides contributes to maintaining soil fertility (Badgley *et.al*., 2007). Organic agriculture is also particularly supported by the CAP (2023-2027)[1] that aims at converting at least 25% of European agricultural areas to organic production within 2030. The importance of having timely and reliable statistical data on the characteristics of organic farms and their related plant and livestock production is evident. This need is further underlined by the entry into force of the SAIO Regulation (EU) 2022/2379[2] (European Union Commission, 2022), which requires the availability of a greater number of statistical data on organic agriculture than in the past. The data on organic farming

[1] The CAP is the Common Agricultural Policy, that rules the financial contributions to the EU farmers.
[2] https://www.europeansources.info/eso_tax_series_titles/official-journal-of-the-european-union/.

disseminated annually by EUROSTAT started in 2013 and is currently updated to 2022[3]. The data collection is ruled by the ESS Agreement on organic production statistics[4]. The main methodological features of organic farming data collection are given in EUROSTAT (2023). In 2022, there were over 402 thousand organic farms in the EU, 58,9% more than in 2013. More than one organic farm in five is Italian (20,5%) and, in Italy, organic farms have grown by 79,8%. Of the approximately 16,5 million hectares of organic Utilized Agriculture Area (UAA) in the EU27, 14,3% are in Italy. Organic surfaces occupy 9,1% of the European agricultural surface, with notable differences among member states. In Italy, the ratio is quite high: 16,8%. In this work, attention is paid to the analysis of the official statistical sources currently existing in Italy (section 2) as regards organic agriculture. A comparison is proposed in section 3 that describes some potential causes of discrepancies. The discrepancies have been analysed using a *micro* approach as well (section 4), while section 5 analyses the odds ratios derived from a logistic model. Some prospective conclusions are drawn in section 6.

2. Main data sources on organic farming in Italy

2.1 The Ministry of Agriculture administrative data

The bodies that carry out checks and certify organic production are authorized by MASAF (Ministry of Agriculture, Food Sovereignty and Forestry). SINAB is the National Information System on Organic Agriculture created by MASAF. SINAB disseminates the "Bio in figures" report every year that includes data on the number of organic operators and organic surfaces by type of cultivation. No data on organic production are available. Data on organic farming distinguish between surfaces already "converted" to organic or still "in the process of being converted". In general, organic farming statistics include both types. The MASAF source is used for providing the yearly data requested by the ESSC 2020/42/6/EN[5], which is going to be replaced by the SAIO Regulation. Access to MASAF microdata is not possible now, because: 1) the farmers provided data for administrative reasons and their use for statistical purposes may conflict with the original arrangements; 2) confidentiality problems.

2.2 The ISTAT structural surveys

The seventh general census of agriculture found its regulatory basis, at the European level, in Regulation (EU) 2018/1091[6]. The agriculture census was mandatory and was carried out by

[3] https://ec.europa.eu/eurostat/data/database.
[4] ESSC 2020/42/6/EN – 12 February 2020.
[5] Data referred to the year Y must be sent to EUROSTAT within June of the year (Y+1).
[6] https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:32018R1091&from=FI.

all the European Union (EU) Member States. To simplify the farmers' response process, the census questionnaire, referring to the date of 1 October 2020, did not ask farmers to indicate the use of the organic practice for each plant species. The main question only asked how many hectares are cultivated using organic farming. Consequently, data relating to the hectares of surface operated with organic practices for each plant species were estimated. Between the 2 censuses of 2010 and 2020, ISTAT measured the evolution of organic farming based on sample surveys conducted in 2013 and 2016 (Integrated Farm Statistics Surveys - IFS). Of course, estimations derived from the IFS sample surveys are affected by the sampling error, which is another reason that may explain discrepancies with the administrative data.

2.3 The FADN survey

The Farm Accountancy Data Network (FADN) is an annual sample survey established by the European Economic Commission in 1965. The FADN field of observation excludes a range of farms based on their economic size. In Italy, the minimum threshold for inclusion is 8.000 euros of annual standard output. The survey is managed by CREA (CREA, 2023), a research body that belongs to the National Statistical System. The survey provides data both on agricultural surfaces and yield per hectare for farms that use organic farming. The last indicator will be used in section 6 to estimate organic production, as requested by SAIO Regulation.

3. Comparison between administrative and survey macro-data in Italy

The MASAF data referred to 2020 are available in SINAB (2021). ISTAT data on organic farming are available online[7]. Table 1 shows the differences between the two data sources as regards 2010, 2013, 2016, and 2020. Except for 2016, survey estimates are always larger than the administrative data: in 2020, the difference was 11,3%, a bit lower than the difference found with reference to the census of 2010 (12,4%). For the whole EU the difference is much higher: 40,1%. While in the North survey data are quite always larger than administrative data, a more uncertain situation characterizes the Centre and the South. In the Islands (Sicily and Sardinia), there is a trend opposite to that of the North. The discrepancies may derive both from different measurements of the organic surfaces given the number of organic farms and from a different number of farms identified as organic by the two sources. In 2020, the MASAF and the ISTAT sources estimated a very similar number of organic farms: 76.449 and 76.084 respectively. Only the Centre area showed a more significant difference in favour of the administrative source (20.552 organic farms versus 15.212). Therefore, discrepancies depend above all on

[7] https://esploradati.istat.it/databrowser/#/.

measurements of organic surfaces, much less on the counting of organic farms. The comparison was conducted for 13 crop species as well (Table 2). Although overall the organic surfaces estimated by ISTAT are higher than those by MASAF, this only occurs for 5 species out of 13. Among the species with large surfaces, the “Other” type shows the highest difference: 650 million hectares from MASAF versus 927 million from ISTAT. The “Other” item includes a wide variety of crops for which the organic feature may be difficult to be certified.

Table 1: Agricultural surfaces managed with organic farming using ISTAT or MASAF data, percent and absolute differences by geographical areas – The years 2010, 2013, 2016, and 2020

GEOGRAPHICAL AREA	ISTAT vs MASAF %				ISTAT vs MASAF			
	2010	2013	2016	2020	2010	2013	2016	2020
Italy (000 ha)	**1.252**	**1.356**	**1.566**	**2.331**	**1.114**	**1.317**	**1.796**	**2.095**
Italy	**+12,4**	**+3,0**	**-12,8**	**+11,3**	**+138**	**+39**	**-231**	**+236**
Northwest	+41,5	+2,8	-24,0	+51,4	+22	+2	-22	+56
Northeast	+31,0	+25,3	+2,6	+41,4	+33	+28	+4	+108
Centre	-5,9	+12,8	-14,0	+7,7	-16	+37	-53	+38
South	+30,9	-2,0	-20,1	+10,8	+108	-9	-131	+75
Islands	-2,6	-4,4	-5,7	-7,7	-9	-19	-29	-41

Source: elaboration on ISTAT and MASAF data.

Table 2: Estimation of organic agricultural surfaces using MASAF or ISTAT data – The year 2020

SPECIES	Agricultural surfaces (ha)		Average % weight (1)	ISTAT vs MASAF %	ISTAT vs MASAF
	MASAF	ISTAT			
Cereals	333.563	354.003	15,5	+6,1	+20.440
Protein cultures	47.057	68.284	2,6	+45,1	+21.227
Root plants	3.494	6.361	0,2	+82,0	+2.867
Industrial crops	43.076	42.127	1,9	-2,2	-949
Fodder crops	426.885	430.124	19,4	+0,8	+3.239
Other arable land	29.471	16.722	1,0	-43,3	-12.749
Vegetables	69.070	37.799	2,4	-45,3	-31.271
Fruit	39.119	37.798	1,7	-3,4	-1.321
Nuts	53.097	47.449	2,3	-10,6	-5.648
Citrus fruits	35.517	30.376	1,5	-14,5	-5.141
Vineyards	117.378	108.345	5,1	-7,7	-9.033
Olives	246.503	224.826	10,6	-8,8	-21.677
Other	650.377	926.880	35,6	+42,5	+276.503
TOTAL	**2.094.607**	**2.331.092**	**100,0**	**+11,3**	**+236.485**

(1) The weight is based on MASAF and ISTAT data. *Source*: elaboration on ISTAT and MASAF data.

We can mention five causes of discrepancies between the two sources. 1) Farms falling below a series of census entry thresholds were not observed. Thresholds may explain cases for which the census provided estimates lower than MASAF. 2) While the administrative source refers to the situation as of 31 December, the census measurement refers to 1 October. 3) The census questionnaire asked farms to indicate whether they practiced organic farming overall. This approximation may have influenced the data relating to the organic surfaces by plant species. 4) The respondent may declare a situation not consistent with the administrative process. For instance, when the certification is at the beginning. 5) The census was carried out during the COVID-19 pandemic when many farms were not active and often the managers were not found. In some periods, direct interviews were even prohibited by the national law.

4. Comparison between administrative and survey data at the micro level

The comparison between the ISTAT and the MASAF sources aims to assess the main features that characterize discordant farms. Since MASAF microdata was not available, we used the microdata supplied by the Italian Integrated Administrative and Control Authority (IACS), which match exactly with the MASAF data. Record linkage was based on the key variable given by the Unique Farm Code. Linked farms were 957.650. Among them, 46.026 records concerned organic farms in both sources and 859,157 non-organic farms in both sources (Table 3). Overall, concordant cases were 94,5%. On the other hand, 52.467 records concerned discordant classifications (the remaining 5,5%): in three out of four cases, discordances concerned farms organic according to the Census and not organic according to MASAF. Non-concordant measurements are much more frequent for farms in the Northeast (8.4% compared to the national average of 5.5%) and, to a lesser extent, in the Centre (5.9%). Other features influence the percentage of concordant cases. Discordant farms are larger. Their average size is 26,6 hectares, while the overall average is 11,9. Moreover, their average number of adult livestock units (ALU) is 12,8, against the national average of 9,2. Their average annual working units (AWU) and standard output are equal to 1,5 and 105 thousand euros respectively, against the national averages of 0,8 and 54 thousand euros. Discordant farms have a manager younger than those that agree (54,5 years versus 61,4). Furthermore, their farm manager has a low educational qualification (up to a middle school diploma) in only 39,1% of cases, compared to 58,7% of cases for the compliant farms. 26,5% of the discordant farms introduced innovations in 2018-2020, versus 11,8% of the concordant ones. Similarly, 37,5% of the discordant farms are digitalized, with only 17% among the concordant ones. Finally, 15,7% of the discordant farms have other gainful activities beyond agriculture, while this share is 6,1% among the concordant ones.

Table 3: Concordance between the MASAF and the ISTAT sources as regards the organic farming practice. Number of farms, the year 2020

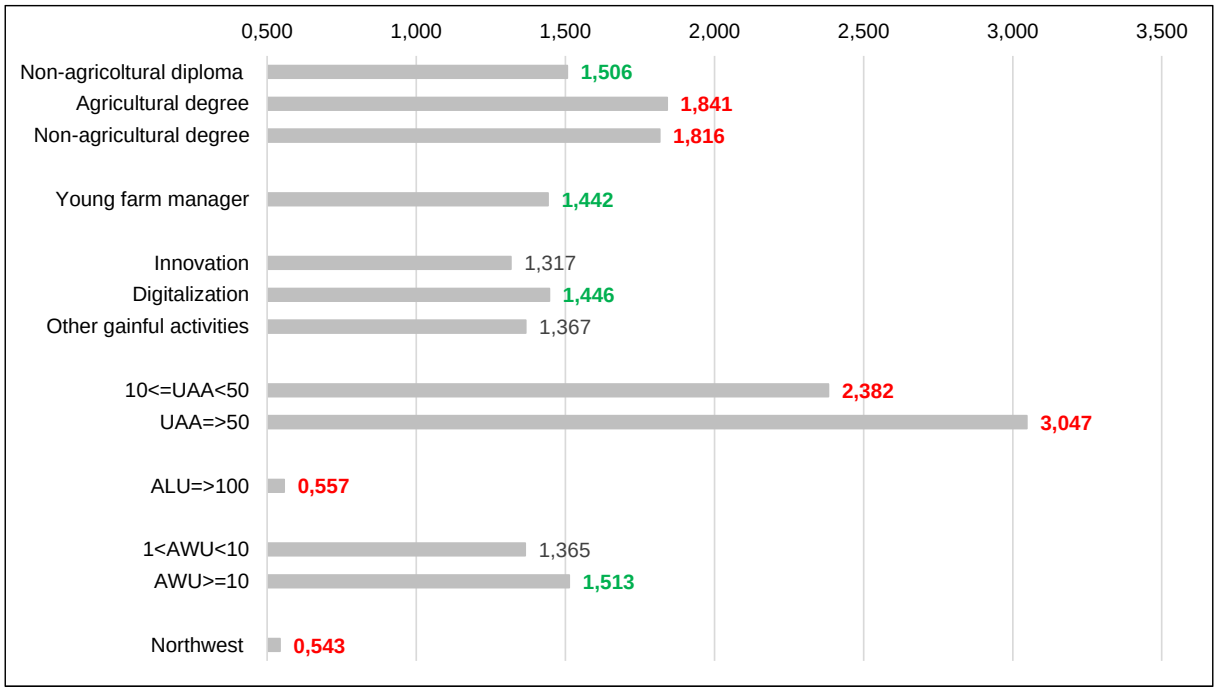
	ISTAT	ABSOLUTE FIGURES			% FIGURES (TOTAL = 100)		
		Organic	Not organic	Total	Organic	Not organic	Total
MASAF	**Organic**	46.026	13.103	59.129	4,81	1,37	6,18
	Not organic	39.364	859.157	898.521	4,11	89,72	93,82
	Total	85.390	872.260	957.650	8,92	91,08	100,00

Source: elaboration on ISTAT and IACS data.

5. The logistic model

Logistic regression (Hilbe, 2009) is used for modeling the probability of an event (dependent variable *Y*) through a series of explanatory *X*-variables. In this context, the *Y* binary variable is equal to 1 if the farm is classified as organic by one source and as not organic by the other, and is equal to 0 otherwise. If *p* is the probability that *Y* takes the value 1 in the presence of a given vector **X** of explanatory variables, the ratio between the probabilities *p* and (1-*p*) is called *odds ratio* and is equal to 1 if and only if *p*=0,5. The odds ratios that are more different from one identify the farms' features that influence more the probability that a farm is discordant. The odds ratios are shown in Figure 1. Results confirm what already seen in section 4.

Figure 1: Odds ratios derived from the logistic model for the most significant dependent variables



Source: elaboration on ISTAT and IACS data.

The most important factor is the agricultural land: farms with at least 50 hectares have a probability to be discordant more than three times larger than those with less than 50. The second most important feature is the degree of education, followed by the number of yearly working units, digitalization and to have young farm manager. Farms in the Northwest and farms with at least 100 adult livestock units have a larger probability of being concordant.

6. Estimation of organic production through data integration

Based on the FADN survey, data on yields were extracted for forty-one varieties, for which estimates referred to 2020 were available for both organic and non-organic farms. The varieties examined represent the 13 plant species listed in Tables 3 and 4. The yield is expressed in quintals of harvested production per cultivated hectare. Although the varieties examined are less than half of those for which the SAIO Regulation requires organic production, the weight of their surfaces was equal to 70%. In nine out of thirteen cases, the organic yield is lower than the yield of non-organic crops. These outcomes confirm the lower average yield of non-organic cultivations (Seufert *et.al.*, 2012; Gismondi, 2022). For each vegetal species, the estimate of organic production is given by the product between the organic Yield derived from FADN and the organic surface derived from the MASAF or the ISTAT sources.

Table 4: Estimation of organic production using MASAF, ISTAT and FADN data – The year 2020

SPECIES	Estimated production (quintals)		Average % weight (1)	ISTAT *vs* MASAF %	ISTAT *vs* MASAF
	MASAF	ISTAT			
Cereals	12.808.819	13.593.714	6,1	6,1	784.895
Protein cultures	2.414.024	3.502.946	1,4	45,1	1.088.922
Root plants	1.014.308	1.846.460	0,7	82,0	832.151
Industrial crops	8.137.056	7.957.762	3,7	-2,2	-179.294
Fodder crops	38.291.585	38.582.101	17,8	0,8	290.517
Other arable land	40.054.036	22.726.843	14,5	-43,3	-17.327.193
Vegetables	79.154.220	43.318.170	28,3	-45,3	-35.836.050
Fruit	8.750.920	8.455.454	4,0	-3,4	-295.466
Nuts	807.074	721.218	0,4	-10,6	-85.857
Citrus fruits	6.524.473	5.579.992	2,8	-14,5	-944.481
Vineyards	12.230.788	11.289.505	5,4	-7,7	-941.282
Olives	6.409.078	5.845.475	2,8	-8,8	-563.603
Other	21.722.592	30.957.787	12,2	42,5	9.235.195
TOTAL (2)	**238.318.974**	**194.377.427**	**100,0**	**-18,4**	**-43.941.547**

(1) Average weight based on both MASAF and ISTAT data. *Source*: elaboration on CREA, ISTAT, MASAF data.

Overall, even if the organic surfaces estimated with the ISTAT source are higher than those estimated with the MASAF source, the opposite occurs with regards to production. This result derives from the fact that, on average, the species for which the MASAF source leads to higher estimated organic surfaces have higher yields than the species for which the opposite occurs.

7. Perspective conclusions

The organic agricultural practice is spreading rapidly, in Italy as well as in Europe (Frascarelli and Baldelli, 2021). As regards 2020, the administrative source produced estimates on the number of organic operators and their agricultural surfaces different from those derived from the agriculture census. The comparative analysis, conducted from both a macro and micro perspective, evaluated the degree of discrepancies and identified their possible causes. Furthermore, we proposed a simple methodology for estimating organic crops production, an indicator requested by SAIO, for which direct measurements are currently not available. Looking ahead, we need to continue comparing the two sources and removing the obstacles that prevent their convergence. At the same time, it is increasingly important to be able to access the administrative database from a micro perspective, with the goal to connect the administrative microdata to the statistical registers. It this way the information content of the administrative data will increase and official statistics on organic farming may be based on administrative data only.

References

Badgley, C., Moghtader J., Quintero E., Zakem E., Chappell M.J., Avilés-Vázquez K., Samulon A., & Perfecto, I. (2007). Organic agriculture and the global food supply. *Renewable Agriculture and Food Systems*, Vol. 22, 2: 86–108.

CREA (2023). *L'indagine RICA*. https://rica.crea.gov.it/.

Eurostat (2023). *Organic farming. Reference metadata in Euro SDMX Metadata Structure (ESMS).* https://ec.europa.eu/eurostat/cache/metadata/en/org_esms.htm.

Frascarelli, A., & Baldelli, M. (2021). Il biologico nella nuova PAC. *Terra e vita*: 21 maggio 2021. https://terraevita.edagricole.it/featured/il-biologico-nella-nuova-pac/.

Gismondi, R. (2022). L'evoluzione dell'agricoltura biologica in Italia: un'analisi basata sull'integrazione tra fonti. *Istat working papers*, 2022/4. https://www.istat.it/it/files//2023/01/IWP-4-2022.pdf.

Hilbe J.M. (2009). *Logistic regression models*. Chapman & Hall. https://www.routledge.com/Logistic-Regression-Models/Hilbe/p/book/9781138106710.

ISTAT (2024). Settimo censimento generale dell'agricoltura: dati definitivi per centro aziendale. https://esploradati.istat.it/databrowser/#/it/censimentoagricoltura/categories/CENSAGR.

Seufert V., Ramankutty, N., & Foley, J.A. (2012). Comparing the Yields of Organic and Conventional Agriculture, *Nature*, Vol.485: 2-29.

SINAB (2021). *Agricoltura biologica: BIO in cifre 2020*. https://sinab.it/sites/default/files/2023-01/230119%20ISMEA%20-%20Bio%20in%20cifre%202021_compressed_19.1.23.pdf.