RAHMANN G & AKSOY U (Eds.) (2014) Proceedings of the 4th ISOFAR Scientific Conference.
'Building Organic Bridges', at the Organic World Congress 2014, 13-15 Oct., Istanbul, Turkey (eprint ID 23618)

Searching for inconsistencies in organic market data – a guide on how to apply quality checks for statistics

CORINNA FELDMANN[1], ULRICH HAMM[1]

Key words: Statistics, organic market data, data quality, data inconsistencies

Abstract

This contribution emerged as part of the collaborative project "Data network for better European organic market information" carried out in the 7th Framework Programme of the EU. To overcome current inconsistencies in organic market statistics in Europe, we generated and applied plausibility checks and equations to organic market data from 39 countries. Thereby we detected inconsistencies, were able to identify their sources and elaborated strategies to harmonise organic market data. We recommend to harmonise data collection and sampling methods, to align nomenclatures and to clearly define product categories to improve organic market data in Europe.

Introduction

Up to now, organic market data collection has been inconsistent throughout European countries; data from different organisations and/or countries is hard to compare, because very different sampling methods, product categories, and nomenclatures have been used. Interpretations based on incomplete and inconsistent data might lead to wrong decisions and misinvestments of companies or policy divisions.

The objective of this contribution is the identification of inconsistencies in organic market data which is currently available throughout Europe. Therefore plausibility checks were applied to data collected through a standardized survey. The survey was led by the Research Institute of Organic Agriculture (FiBL, Switzerland) and the Agricultural Market Information Company (AMI, Germany) (cf. Willer and Schaack 2013). Plausibility checks were conducted one by one for the survey responses from data collectors in 39 countries in Europe.

Material and methods

The quality of organic market data was checked by plausibility equations; some important ones are listed in Figure 1. These plausibility checks have been generated based on the experiences of the project 'Organic Marketing Initiatives and Rural Development' (cf. Hamm and Gronefeld, 2004) and adjusted to the needs of this data analysis. All checks have been carried out with Excel software. For some of the checks formulas were entered in an additional column in Excel and exceptional results were highlighted through the 'conditional formatting' function. The remaining plausibility checks, mainly comparisons between two years or countries with similar production conditions were carried out through a thorough, manual search. The findings were distributed among the partners responsible for the survey execution in each country. Especially those six countries which participate in the project's case studies were asked to investigate the origin and causes of the inconsistencies. The outcome of these investigations will increase the awareness of data collectors and will feed in the upcoming case studies.

[1]University of Kassel - Department of Food and Agricultural Marketing, Germany, e-Mail: c.feldmann@uni-kassel.de, http://www.uni-kassel.de/fb11agrar/fachgebiete-einrichtungen/agrar-und-lebensmittelmarketing/startseite.html

1. Comparison between two years (e.g. 2010-2011) for area, production, and sales data
2. Comparison between countries with similar farming conditions
3. Organic production (share in %) < organic area (share in %)
4. Organic yield < conventional yield
5. Organic area < total area
6. Imports < Sales
7. Domestic organic consumption = organic sales, sold as organic + organic imports – organic exports
8. Market share (volume) = organic consumption / total consumption × 100
9. Import share of organic human consumption = organic imports for human consumption / organic human consumption × 100
10. Export share of organic sales from domestic production = organic exports for human consumption / organic sales as organic for human consumption × 100
11. Degree of self-sufficiency = sales of organic as organic for human consumption / organic human consumption × 100

Figure 1: Plausibility checks to determine inconsistencies and mistakes

Results and conclusions

The extent of the results depends on the details and the amount of available data in the respective country. Due to a lack of organic market data in many countries, more complex equation models to validate data consistency (e.g. domestic organic consumption = organic sales, sold as organic + organic imports – organic exports) could not be applied in most of the countries. Since especially area data was widely available, most inconsistencies were detected in this type of data. Consequently, most inconsistencies were found through comparisons of area data between two years or countries with similar production conditions. Furthermore, a lot of unrealistic values (e.g. negative numbers) were detected, when carefully scanning the data of each respective country. Below some examples for different types of comparisons to reveal inconsistencies are shown:

- Example for a comparison of area data between two years:

  Peas, fresh: 108.2 ha (2010) and 0.8 ha (2011)

- Example for a comparison of production data between two years:

  Pork: 12,540.0 t (2010) and 5,400.0 t (2011)

- Example for a comparison of yields (organic yield < conventional/total yield):

  Strawberries: 5.83 t/ha (organic) > 3.29 t/ha (conventional)

The inconsistencies that could be revealed were grouped according to data type and also according to data origin. Figure 2 is an example for the presentation of the inconsistencies resulting from descriptive statistics. In this figure inconsistencies in organic area data are shown, summed up depending on the status of EU membership. Further figures include yield, production, and sales data, that is analysed on a country specific level (i.e. for all 39 countries, if data was available) and also aggregated into country groups based on the status of EU membership. Thus it is possible to show the number of organic data collected, the number of possible comparisons that could be conducted for each plausibility check, the number of inconsistencies found through the application of these plausibility checks, separately for each country and data type. For area (cf. Figure 2) and sales data, most data is available from EU15 member states and hence the number of possible comparisons is also higher than for the other country groups.

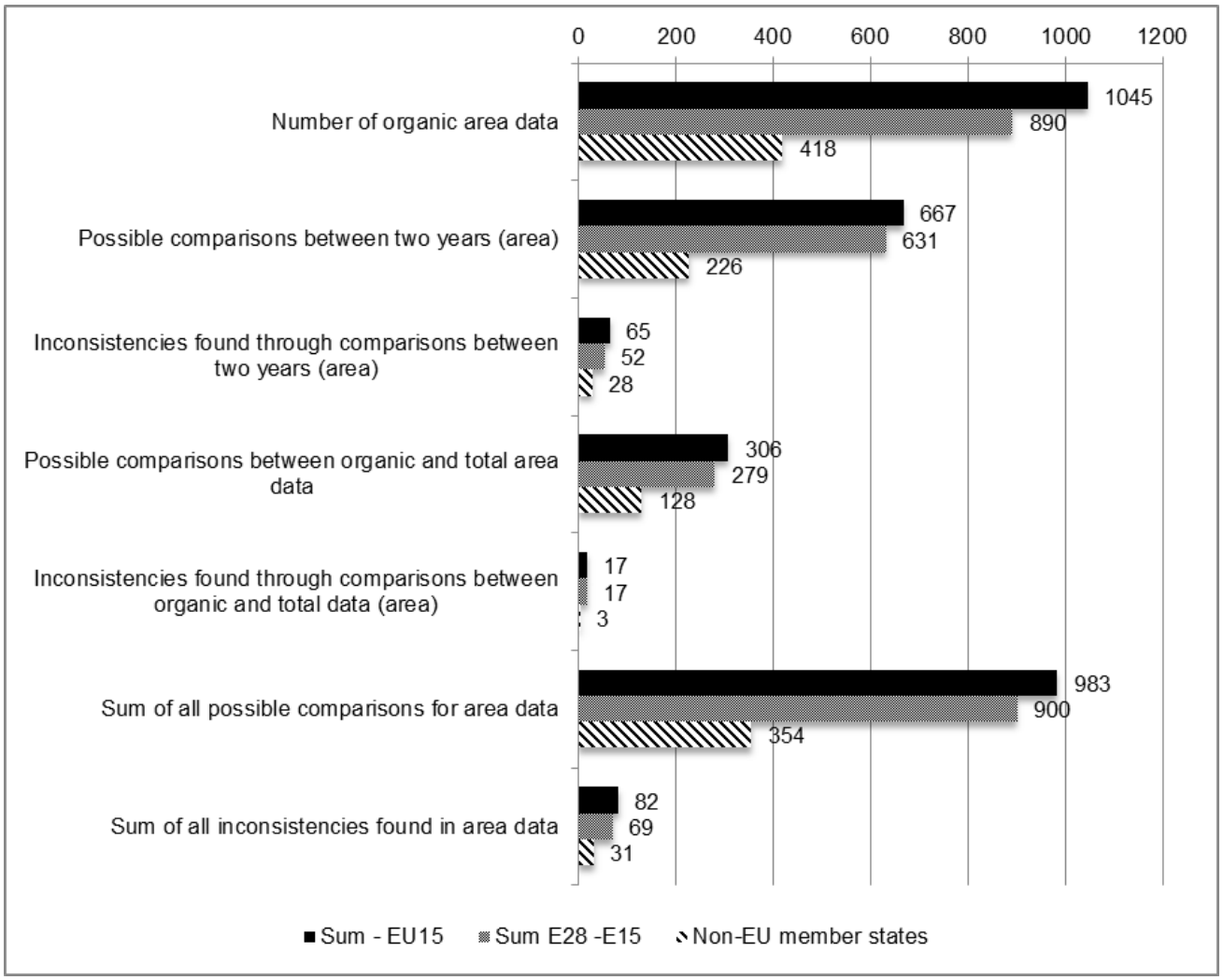

Figure 2: Inconsistencies in organic area data

Concerning organic production and yield data, more data is available from the group of 13 countries, which have joined the EU since May 2004. In this group of countries, especially Bulgaria, Croatia, Czech Republic and Poland delivered a great amount of data (Figure 3).

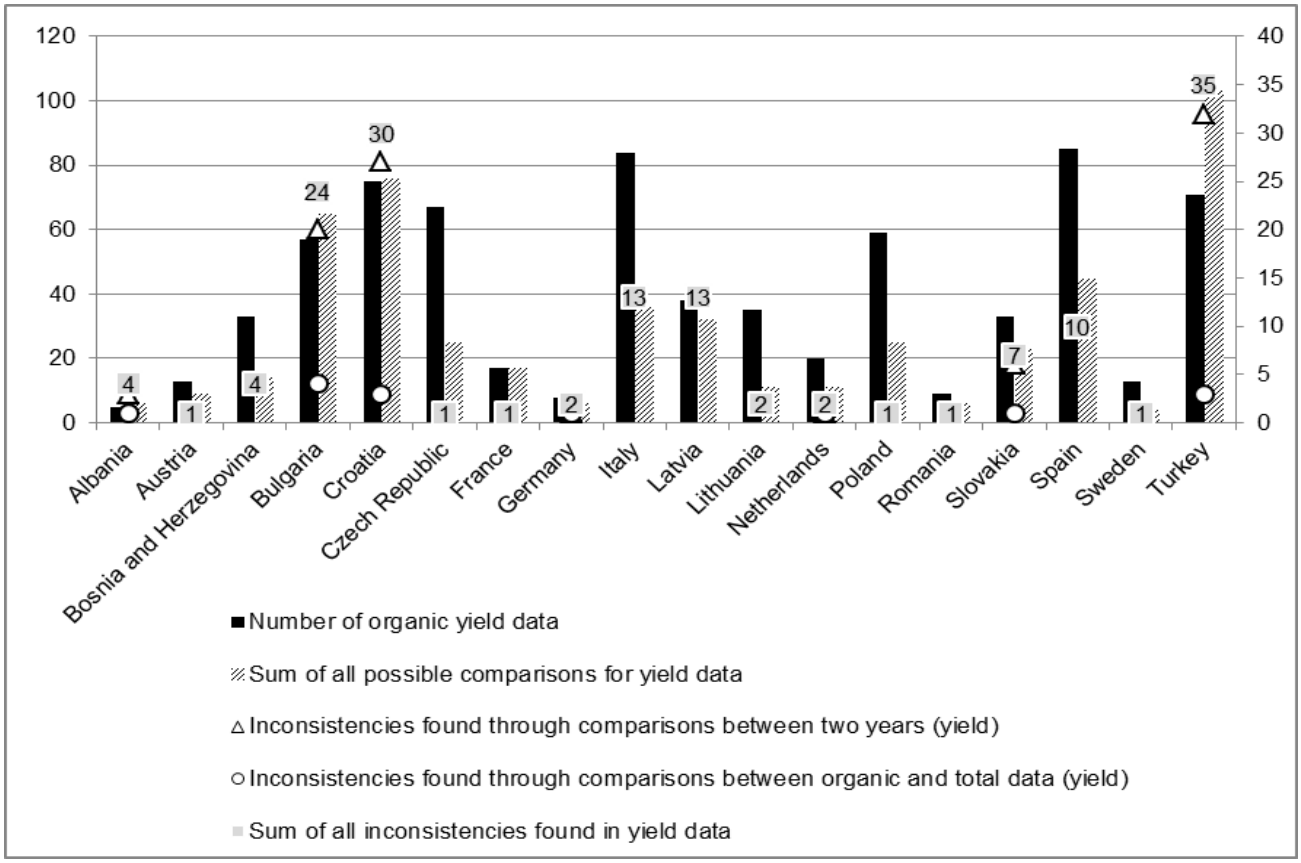

Figure 3: Inconsistencies in organic yield data

Most mistakes in organic market data stem from heterogeneous nomenclature and varying definitions of product categories. Inconsistencies in data from two subsequent years often occur because data from different sources was used. Therefore, it is very important to harmonise data collection methods, product categories, and nomenclature to ensure coherence and comparability. Likewise, inconsistencies might result from the estimation of numbers from organic market experts; these tend to be higher than the actual numbers and lack transparency and traceability.

Discussion

Further steps resulting from the outcome of the data plausibility checks are the compilation of a guideline for organic market data collectors, the exchange of opinions and experiences within the organic data network, and the revision of current organic market data reports. The following issues will be addressed: recommendations for consistent data collection and sampling, ideas for a harmonised nomenclature, clear definitions of product categories, and presentation of the data quality check procedure.

References

Willer, H. and Schaack, D. 2013, Intermediate report on compilation of key organic market data, Deliverable 4.2, Organic Data Network - Project no. 289376, Frick and Bonn.

Hamm, U. and Gronefeld, F. 2004, The European Market for Organic Food: Revised and Updated Analysis, Organic Marketing Initiatives and Rural Development, Vol. 5, University of Wales, Aberystwyth, United Kingdom (ISBN 0-9543279-4-7).