Statistics Lithuania

IMPROVING METHODOLOGICAL PROCESSES TO BE COMPATIBLE WITH THE NEW LEGAL FRAMEWORK OF IFS: INTEGRATION OF EXISTING ADMINISTRATIVE DATA SOURCES; GEOCODING/GEOLOCATION OF FARMS

Eurostat Grant Agreement No 218.0216

FINAL TECHNICAL AND METHODOLOGICAL REPORT
February 2020, Vilnius

Project Manager
Aušra Jablonskienė
Head, Agricultural and Environmental Statistics Division
Statistics Lithuania

Content

1. Introduction

Agricultural statistics are aimed at producing data on agriculture that meet the current and future user needs in an efficient manner with a bearable burden on data producers.

A strategy on agricultural statistics that has been drafted by Eurostat and agreed with the European Statistical System Committee (ESSC) has several specific objectives. One of these objectives is to produce more statistics while lowering the burden on respondents by exploring alternative data sources and possibilities of efficiency improvement.

Statistics Lithuania already uses administrative data for producing agricultural statistics. The following administrative data are used: the Integrated Administration and Control System (IACS), Farm Animal Register, Organic Farming Register. In recent years more and more administrative data are used directly, without questioning farms. Moreover, administrative data are used for prefilling questionnaires as well as for data checking.

However, in recent years the possibility to use a new data source – the Database of the State Social Insurance Fund Board under the Ministry of Social Security and Labour (hereinafter referred to as SSIFB) – has emerged. Some years ago (for example in 2013 and earlier) only a minority of farms registered their non-family labour force in the SSIFB and usage of the SSIFB data was impossible. During the analysis of administrative data sources which could be used for the Agricultural Census 2020 needs, it was clarified that recently farmers, especially owners of large farms, register their non-family labour force in the SSIFB due to the new requirements and new legislation. Therefore, it was very important to analyse this new data source and to find ways how it can be used for production of agricultural statistics. It was very important to compare data from the SSIFB with data collected during statistical surveys. The main survey in agricultural statistics in which data about non-farm labour force are collected is the Farm Structure Survey (hereinafter referred to as FSS). Therefore, it was decided to compare the SSIFB data with data of the FSS 2016.

Moreover, one more new emerging need was detected. The geo-referenced information becomes more and more important and is to be included in as many data collection exercises as possible to allow for data linkage. Data collection exercises in the agricultural domain are not an exception.

The Regulation (EU) 2018/1091 of the European Parliament and of the Council of 18 July 2018 on Integrated Farm Statistics and repealing Regulations (EC) No 1166/2008 and (EU) No 1337/2011 (hereinafter referred to as IFS Regulation) foresees that a Statistical Unit theme grid in accordance with Annex III of Directive 2007/2/EC of the European Parliament and of the Council should be used for defining the geo-reference of agricultural holdings. Therefore, it is necessary to develop the methodology on how to determine the grid for each holding.

2. PROJECT IMPLEMENTATION

2.1. Project objectives

The first objective of this project is to develop and test a new methodological approach towards use of data from the SSIFB for fulfilling the requirements of IFS Regulation, specifically, for the module Labour Force and Other Gainful Activities.

The second objective of the action is to develop the methodology for defining the geographical location of the holding as it is required by the IFS Regulation.

2.2. Project timetable

Milestones/ Deliverables / Tasks / Results / Reports	Timetable
Establishment of a working group and preparation of action plans	M+1
Integration of existing administrative data sources	
Initial analysis of the Database of the State Social Insurance Fund Board under the Ministry of Social Security and Labour (hereinafter referred as SSIFB)	M+2
Analysis of the SSIFB data structure, identification of the indicators which could be used for the needs of IFS Regulation (the module Labour force and Other Gainful Activities)	M+3
Identification of the methodological consistency of indicators in the SSIFB database with indicators of IFS Regulation (the module Labour force and Other Gainful Activities)	M+4
Estimations of characteristics for the implementation of IFS Regulation using the SSIFB data and farms from a sample of the FSS 2016	M+6
Analysis of the results gained during the estimations and their comparison with the FSS 2016 data	M+8
Identification of the SSIFB data which can be used for Agricultural Census 2020, FSS 2023 and FSS 2026	M+9
Development of the Agricultural Census 2020 methodology for the module Labour force and Other Gainful Activities taking into account the experience gained and the results achieved during the action	M+11
Preparation and transmission of the interim technical report (including works done regarding geocoding)	M+12
Geocoding/geolocation of farms	
Studying of methodological materials and experience of other countries on defining the geographical location of the holding	M+5
Analysis of administrative data sources, consultations with administrative data managers on possibility to use administrative georeferenced data for statistical purposes	M+7
Preparation and signing new agreements with administrative data managers (or supplementing the existing ones) on the taking and use of georeferenced data (from the Identification System for Agricultural Parcels in IACS)	M+10
Taking of the administrative georeferenced data, their analysis and preparation for use for agricultural statistics purposes	M+12
Creating the methodology for defining the geographical location of the holdings as it is required by the IFS Regulation for farms registered in the IACS	M+13

Testing the created methodology for defining the geographical location of the holdings which declared their Utilised Agricultural Area in the IACS in 2018	M+15
Creating the methodology for defining the geographical location of the holdings which are not registered in the IACS	M+17
Testing the methodology for defining the geographical location of the holdings which did not declare their Utilised Agricultural Area in the IACS in 2018	M+18
Preparation of the final technical and methodological report	M + 20
Final report on implementation of the action and financial statement	60 days following the closing date of the action

All works foreseen in the project timetable were carried out on time. There were no significant delays during the project, works were carried out smoothly.

2.3. Project implementation

First of all, the list of the staff members – a Working Group – involved in the project was drawn up and approved by the order of the Director General (Order No 184 of 21 August 2018). The Working Group for the project consists of the specialists of the Agricultural and Environmental Statistics Division, Methodology and Quality Division, Statistics Dissemination and Communication Division, European Affairs and International Cooperation Division and Performance and Finance Management Division of Statistics Lithuania. The action plan for each member of the said Working Group was created.

The initial analysis of the Database of the SSIFB was carried out. First of all, the Lithuanian legislation related to labour accounting was analysed in order to understand main accounting principles.

2.3.1. Works related to use of SSIFB data for the Agricultural Census 2020

Analysis of the SSIFB data structure, identification of the indicators which could be used for the needs of IFS Regulation (the module Labour force and Other Gainful Activities) was carried out as well. Statistics Lithuania has an agreement with the SSIFB for getting the data from the SSIFB database. The last update of this agreement was done in April 2019. During the analysis of the SSIFB data structure, the list of farms which were involved in the Farm Structure Survey (hereinafter referred to as FSS) 2016 sample was taken and data from the SSIFB were linked to these farms (at farm level). Non-family labour force was analysed. This was done separately for farmers' and family farms (natural persons) and agricultural companies and enterprises (legal persons). A short overview of this analysis will be shown in the next section of this report.

In order to understand if data received from the SSIFB could be used for the FSS needs, the methodological principles of creation of the SSIFB database were analysed. The methodological consistency of indicators in the SSIFB database with indicators of IFS Regulation was analysed as well. The main inconsistency was found out: in the SSIFB, employees are registered without indicating how much time they work, while in the FSS and Agricultural Census it is required to indicate the number of persons (non-family labour force, directly employed by the farm on the regular

basis) per AWU band (0-<25, ≥25-<50, ≥50-<75, ≥75-<100, (100)). It means that employers register their employee in the SSIFB, but there is no requirement to indicate how much time this employee works per day (or per month).

After linking the SSIFB data to the FSS 2016 data, estimations of characteristics for the implementation of IFS Regulation using the SSIFB data and farms from a sample of the FSS 2016 were made and analysis of the results gained during these estimations and their comparison with the FSS 2016 data was performed. The main outcome of this analysis will be shown in the next section of this report.

The SSIFB data which can be used for the Agricultural Census 2020, FSS 2023 and FSS 2026 were identified and the main principles of the Agricultural Census 2020 for the module Labour Force and Other Gainful Activities were developed.

The Agricultural Census 2020 methodology for the module Labour force and Other Gainful Activities was developed, taking into account the experience gained and the results achieved during the project.

2.3.2. Works related to geocoding of farms

In parallel with the works described above, works related to geocoding farms were carried out as well. Methodological materials and experience of other countries on defining the geographical location of the holding were studied. Several data sources, which could be used for statistical and geospatial data integration, were analysed. Available information and experience from other statistical divisions of Statistics Lithuania in using geospatial information was analysed. Information from the State Enterprise Centre of Registers (Address Register) was studied, and it was found out that this information could be used for production of agricultural statistics as additional data source.

However, consultations with administrative data managers on possibility to use administrative georeferenced data for statistical purposes have shown that the best solution is to use the georeferenced data from the Identification System for Agricultural Parcels in IACS.

Due to the fact that Statistics Lithuania has not used these georeferenced data so far, an official letter was prepared and sent to the State Enterprise Agricultural Information and Rural Business Centre (hereinafter referred to as AIRBC) in November 2018 (Official Letter of 26 November 2018 No SD-861) with the justification for the said data. However, such data had not been received yet, AIRBC asked for more clarification and justification. Therefore, the second official letter was prepared and sent to AIRBC (Official Letter of 12 December 2018 No SD-950). When all uncertainties had been clarified, AIRBC prepared and delivered to Statistics Lithuania digital data of plots of agricultural land and crop fields declared in IACS in 2018.

Data were formed so as to be possible to identify an applicant (holding) and to link his data with the data from other registers (Address and Population Register, Land Cadaster). It was agreed that the nature of data provision will be one-time. It means that no new agreements will be signed with AIRBC (or the existing ones supplemented) and data for the Agricultural Census 2020 needs will be received according to the special request (the official letter), i.e. in the same way as they were received for the needs of this project.

Georeferenced data were analysed. The results of this analysis will be shown in the next section of the present report.

After analysis of the georeferenced data, the geographical location of the farms registered in the IACS was defined. The methodology for defining the geographical location of the holdings as it is required by the IFS Regulation for farms registered in the IACS was developed. For testing the methodology, the georeferenced data on Utilised Agricultural Area declared in the IACS in 2018 were used.

However, not all farms declared their Utilised Agricultural Area in the IACS. Therefore, the methodology on how to define the geographical location of the holdings which do not declare their Utilised Agricultural Area was necessary to develop. And it was done using other data sources, such as the Real Estate Register, Address Register, Agricultural Census 2010 and FSS 2013 data. The way how to define the geographical location of the holdings which do not declare their Utilised Agricultural Area in IACS was found and methodology for this defining was developed.

3. RESULTS OF THE ANALYSIS AND THE METHODOLOGIES CREATED

3.1. USE OF THE DATA FROM THE DATABASE OF THE SSIFB FOR FULFILLING THE REQUIREMENTS OF IFS REGULATION

3.1.1. Short information on the data collection process in the FSS 2016

Farmers' and family farms
During the FSS 2016, data collection process (survey fieldwork) was subcontracted, i.e. services of a market research company were purchased. The market research company collected data from farmers' and family farms.
Farmers had a possibility to fill in and transmit the filled-in questionnaires to Statistics Lithuania themselves via the Internet using a new electronic statistical data preparation and transfer system e-Statistics for the Population. The market research company collected data from those farmers who had not filled in the questionnaires themselves.

All farmers and family farms selected for the Farm Structure Survey 2016 were divided into two groups. One group covered farms with 10 or less hectares of utilised agricultural land (about 15 000 farms); these farms were surveyed by phone. The remaining farms (about 30 000 farms with utilised agricultural land of more than 10 hectares) were included into the other group; these farms were surveyed face-to-face.

Agricultural companies and enterprises

Agricultural companies and enterprises filled in the questionnaire themselves and sent it directly to Statistics Lithuania for further processing via the electronic statistical data preparation and transfer system e-Statistics.

3.1.2. Analysis of the Database of the SSIFB

First of all, a list of farms which were involved in the FSS 2016 sample was taken and data from the SSIFB were linked to these farms (at the farm level). SSIFB data allow analysis of non-family labour force only because family labour force is not registered in this database.

Non-family labour force (directly employed by the farm on the regular basis) of farmers' and family farms (natural persons) and agricultural companies and enterprises (legal persons) was analysed separately.

Analysis of the non-family labour force in farmers' and family farms

During the analysis, data of 45 011 farmers' and family farms were linked (FSS 2016 and SSIFB data). It is important to mention that only those farms which are registered in the Farmers' Farm

Register and have Farmers' farm identification number can register data in the SSIFB database. It was found out that out of 45 011 farms 9 154 (20 per cent) were not registered in the above mentioned register. Therefore, it was possible to analyse data only of 35 857 farms.

According to the SSIFB data, only 4 756 (13 per cent) farms out of 35 857 have non-family labour force (according to the FSS 2016 data – 4 242 (12 per cent)). In 1 730 farms, data from the SSIFB and from the FSS 2016 are equal; in 1 179 farms, the number of non-family labour force is higher in FSS 2016 than in SSIFB; in 2 357 farms, the number of non-family labour force is higher in SSIFB than in FSS.

It was useful to analyse such differences more deeply. In Table 1, the number of farmers' and family farms in which differences in the SSIFB and the FSS 2016 data were found is distributed according to the size of difference.

Table 1. Distribution of farmers' and family farms by size of differences (the SSIFB data compared to the FSS 2016 data)

Difference by number of non-family labour force	Number of farms
1	2 246
2–4	1 109
5–9	127
10 and more	52
Total	**3 534**

Figure 1. Distribution of farmers' and family farms by size of differences (the SSIFB data compared to the FSS 2016 data), per cent
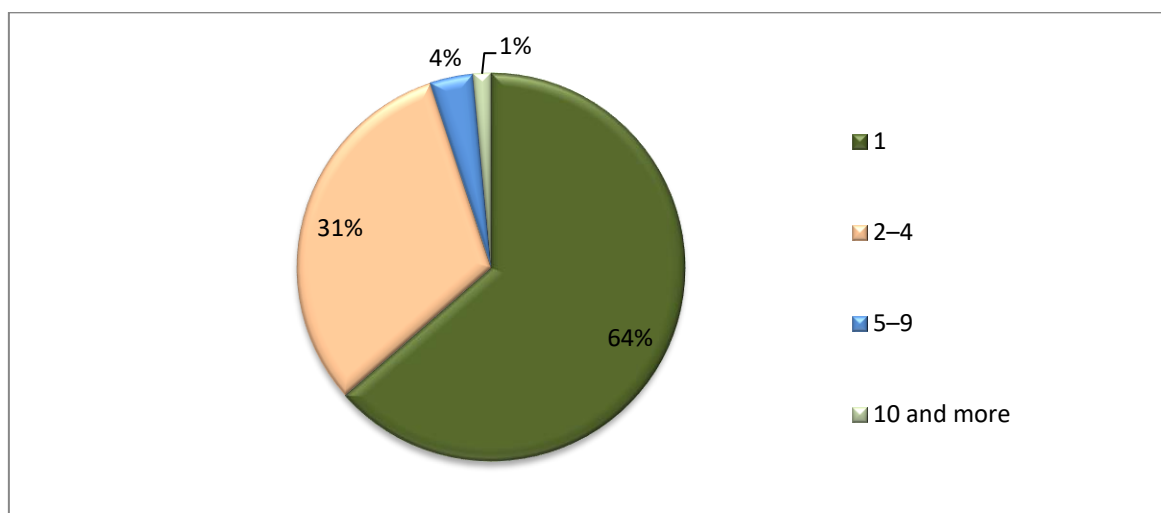
We can see that in most farms (64 per cent) the difference is 1 employee. However, there were 179 farms where the differences were more significant (5 or more employees). Difference of 2–4 can also be considered significant.

After the analysis of all the differences, further estimations were done. The estimation of non-family labour force in farmers' and family farms from the FSS sample which are registered in the Farmers' Farm Register and have the farmers' farm identification number (35 857 farms) is presented in Table 2.

Table 2. Non-family labour force in farmers' and family farms

	FSS 2016	SSIFB	FSS 2016 compared to SSIFB, per cent
Non-family labour force (male)	11 249	12 209	92
Non-family labour force (female)	4 647	6 444	72
Total	**15 896**	**18 653**	**85**

Figure 2. Non-family labour force in farmers' and family farms
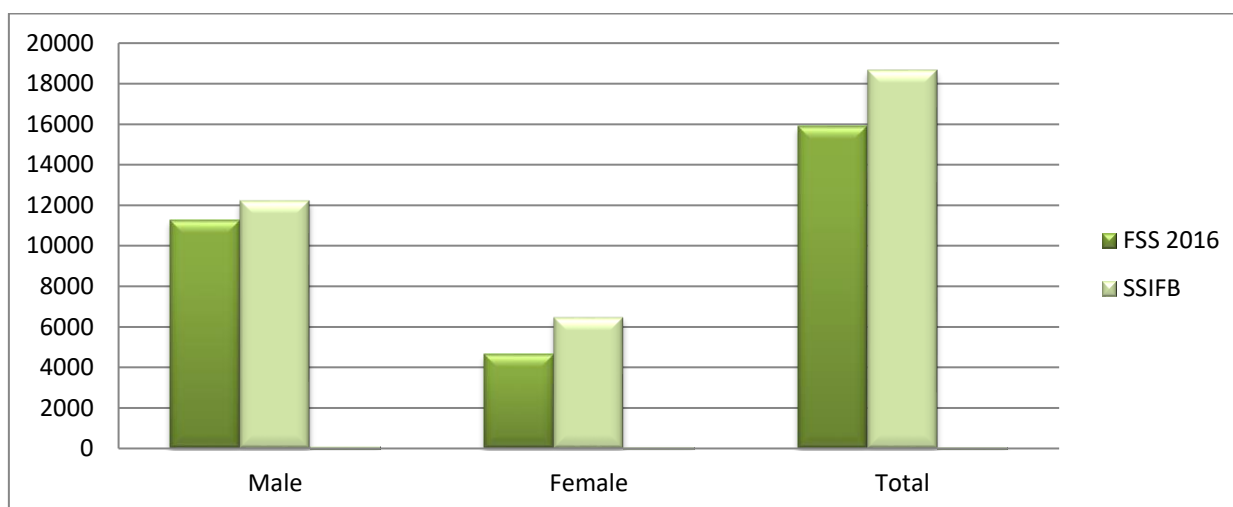
We can see that the total number of non-family labour force according to the FSS 2016 data is smaller than in the SSIFB database, and the difference is quite significant. Therefore, the use of the SSIFB data directly may lead to significant inconsistency.

Analysis of the non-family labour force in agricultural companies and enterprises

During the analysis, data of 896 agricultural companies and enterprises were linked (the FSS 2016 and the SSIFB data).

In 194 enterprises, data from the SSIFB database coincide with the data of the FSS 2016. In 27 enterprises, the number of labour force in the FSS 2016 is higher than in the SSIFB database; in 675 enterprises, the number of labour force in the SSIFB database is higher than in the FSS 2016. It is clear that in the majority (75 per cent) of enterprises the number of labour force in SSIFB is higher than in FSS 2016. Main reason of such phenomenon is that in most of the enterprises not only employees who carry out farm work, but also employees who carry out work other than farm work are employed. It is important to mention that the differences in the SSIFB and the FSS 2016 data are significant (Table 3). In 45 enterprises, the difference is more than 10 employees.

Table 3. Distribution of agricultural companies and enterprises by size of differences (SSIFB data compared to FSS 2016 data)

Difference by number of labour force	Number of enterprises
1	128

2–4	152
5–9	108
10 and more	314
Total	**702**

Figure 3. Distribution of agricultural companies and enterprises by size of differences (SSIFB data compared to FSS 2016 data), per cent
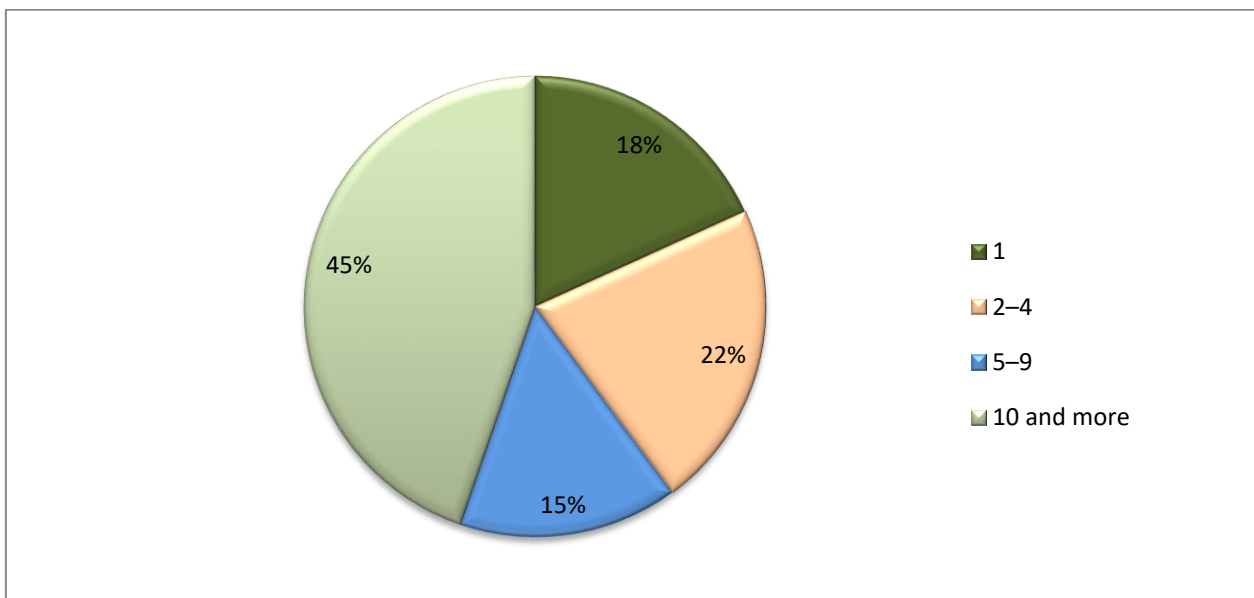
Respectively, significant differences were determined during the estimation of the total number in agricultural companies and enterprises (Table 4).

Table 4. Labour force in agricultural companies and enterprises

	FSS 2016	SSIFB	FSS 2016 compared to SSIFB, per cent
Labour force (male)	10 743	24 504	44
Labour force (female)	6 597	19 565	34
Total	**17 340**	**44 069**	**39**

Figure 4. Labour force in agricultural companies and enterprises
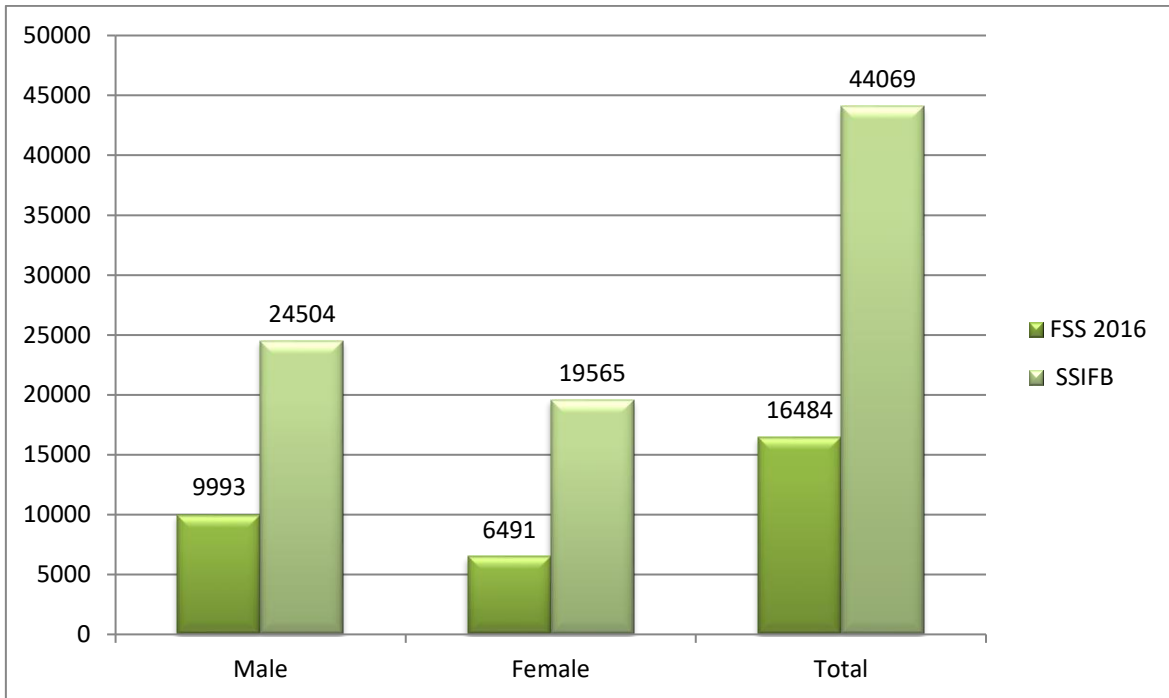
3.1.3. Main principles on the use of SSIFB data during the Agricultural Census 2020 and FSS 2023, 2026

1) Variables about non-family labour force should be included in the Agricultural Census 2020 (as well as in the FSS 2023 and 2026) questionnaires: in the questionnaire for farmers' and family farms as well as for agricultural companies and enterprises.

2) The SSIFB data cannot be used directly, without questioning.

These two statements are illustrated in Figure 5. The differences in SSIFB and FSS 2016 data are noticeable.

Figure 5. Non-family labour force
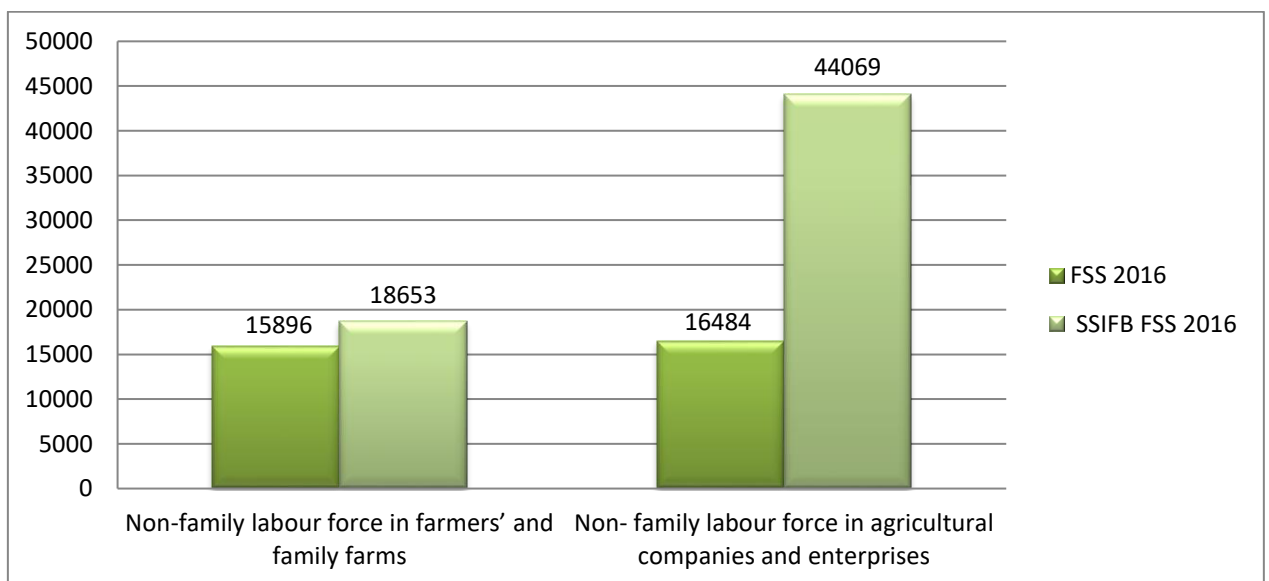
3) A number of non-family labour force (total and distributed by gender) can be prefilled in the farmers' and family farms questionnaire. However, a possibility to correct the data if they are inaccurate should be left for farmers. Moreover, there is no distribution of labour force by working hours in the database of SSIFB. Therefore, not all data needed for the Agricultural Census 2020 and for the FSS 2023 and 2026 may be found in the database of SSIFB.

4) There is no possibility to use the SSIFB data for data collection from agricultural companies and enterprises. The analysis showed that the SSIFB data and the FSS 2016 data differ significantly. In the SSIFB database there is no possibility to separate employees who carry out farm work and who carry out work not related to farm work in companies and enterprises the main activity of which is not agriculture. Moreover, companies and enterprises the main activity of which is agriculture may carry out other activities not related to agriculture. Therefore, the prefilling of the SSIFB data in the questionnaire for agricultural companies and enterprises may cause respondents' dissatisfaction and may lead to preparation of incorrect statistical information.

3.1.4. Methodology of the Agricultural Census 2020 for the module Labour force and Other Gainful Activities

During the Agricultural Census 2020 as well as IFS 2023 and 2026, Lithuania will collect statistical data about farm labour force as it is foreseen in the IFS Regulation. 'Labour force and Other Gainful Activities' module variables will be included in the Agricultural Census 2020, IFS 2023 and 2026 questionnaires. All variables required by the EU legal acts will be included. There is no possibility to remove one or several variables from the above mentioned questionnaires due to inconsistencies in the FSS 2016 and SSIFB data which were found during this project.

Two separate questionnaires will be prepared for each statistical survey (Agricultural Census 2020, IFS 2023, 2026): first one – for farmers' and family farms, second one – for agricultural companies and enterprises.

Questionnaires for agricultural companies and enterprises will not be prefilled.

Only questionnaires for farmers' and family farms will be prefilled with data form the SSIFB database. Two variables will be prefilled: the total number of non-family labour force (male) regularly working on the farm and total number of non-family labour force (female) regularly working on the farm, i.e. the last column of the following table (Table 5) will be prefilled.

Table 5. Table in the Agricultural Census 2020, IFS 2023 and 2026 questionnaires
Number non-family labour force

	Code	Average number of hours worked per day					Total number of workers (sum of columns 1–5)
		Under 2	2–4	4–6	6–8	8 and over	
A	B	1	2	3	4	5	**6**
Male	581						
Female	582						

During the data collection, the farmers will have the possibility to correct the data if they are inaccurate. It means that if farm has more (or less) workers than it is prefilled, farmer may delete incorrect data and write correct ones.

However, farmers will have to distribute the total number of non-family labour force (male/female) by average number of hours worked per day, because there is no distribution of labour force by working hours in the database of SSIFB.

The order of getting the SSIFB data at Statistics Lithuania

Statistics Lithuania receives the SSIFB data according to the agreement. The owner of these data in Statistics Lithuania is the Enterprise Statistics Division. This division is the main user of the SSIFB data.

First of all, the Agricultural and Environmental statistics division of Statistics Lithuania fill in the application for getting data from the SSIFB. According to the rules for protection of personal data and rules of confidentiality at Statistics Lithuania, such application is needed for getting any administrative data from other divisions of Statistics Lithuania.

The application will be submitted to the Enterprise Statistics Division on May 2020, 2023 and 2026. The reference day for the Agricultural Census 2020 and IFS 2023 and 2026 is 1 June 2020, 2023, 2026 accordingly. Therefore, data about persons who worked on farms during 12-month period ending on a reference day should be prepared.

After receiving the above mentioned application, the Enterprise Statistics Division forwards it to the IT Division. After that, by the end of June, the IT Division prepares the SSIFB data for the Agricultural and Environmental Statistics Division. Data are prepared at farm level, each worker is recorded separately.

Specialists of the Agricultural and Environmental Statistics Division analyse these data, do necessary estimations and send these data to IT specialist responsible for programming survey e-questionnaire for prefilling.

The application contains the following information:

1. Data description. Data about each worker, who worked on farm during last 12-month period ending on a reference day: ID code, sex, date when worker started their work on farm, date when worker finished their work on farm (if finished);

2. Purpose of data use (why data is needed);

3. Legal base for getting the data;

4. Criteria for data presentation;

5. Data transmission way;

6. Data format;

7. The date of destruction of administrative data;

8. Specialist of the Agricultural and Environmental Statistics Division who will work with the administrative data (the SSIFB data).

The task for IT specialists

Questionnaires of each survey are discussed with IT specialists. After that the task for primary data entry is prepared for IT specialists.

In the case of the Agricultural Census 2020 questionnaire, the task for IT specialists also should be prepared. The condition that the Agricultural Census 2020 questionnaire should be prefilled with SSIFB data is included in this document.

3.2. DEFINING THE GEOGRAPHICAL LOCATION OF THE HOLDING AS IT IS REQUIRED BY THE IFS REGULATION

3.2.1. Order of works carried out

The main task for specialists involved in this project was to analyse the possible data sources and determine how they can be used for defining the geo-reference of agricultural holdings.

According to the tasks involved in the time table of the project, two pieces of works were foreseen. One piece was related to agricultural holdings, registered in IACS, another one – related to agricultural holdings, not registered in IACS. Therefore, works were organized so, that both pieces of works would be carried out fully and the results gained would be used for the Agricultural Census 2020.

Agricultural holdings registered in IACS

On 27 December 2018, the AIRBC prepared and delivered to Statistics Lithuania digital data of plots of agricultural land and crop fields declared in IACS in 2018. Data were received in *shapefile* format.

All steps related to the digital data received from the AIRBC were made using ESRI ArcGIS software. Geospatial data is in the LKS-94 coordinate system.

Digital data of more than 1 million fields (1 148 900 fields) of agricultural areas were received from the AIRBC. These fields were linked to 128 142 different owners of farms (agricultural holdings). By using ID code of farm owners, the above mentioned 128 142 farms were linked to the Statistical Farm Register.

After the update of the Statistical Farm Register, it was found out that 163 693 farms were registered in this register in the beginning of 2019.

A successful linkage was done for 123 361 farms. These farms comprise 75 per cent of all the farms in the Statistical Farm Register. It means that for these farms geocoding will be done without any additional steps, just using digital data from the AIRBC.

4 781 farms from the AIRBC were not found in the Statistical Farm Register. These farms were analysed separately. 968 of these farms are agricultural companies and enterprises, they are not included in the Statistical Farm Register, and they are included in the Business Register. Also, part of these farms has less than 1 ha utilised agricultural land or has only such areas which are not agricultural land, such as Nature 2020, forest land, etc. Also, there were farms in which farm holder has changed. In fact, some holdings exist in the Statistical Farm Register, but with an old farm holder, because the old farm holder was in the Statistical Farm Register at the time of its update.

After this analysis it was decided not to further analyse 3 813 farms, i.e. those farms which have less than 1 ha utilised agricultural land, which have only such areas which are not agricultural land (such as Nature 2020, forest land, etc.), and which already exist in the Statistical Farm Register, but with an old farm holder (because they are duplicates and only the farm holder will be updated in the Statistical Farm Register, the new farm will not be included).

After all actions described above, plots of all fields of 123 361 farms were estimated and the largest field of each farm was defined. Moreover, the centre point coordinates of these largest plots were defined.

Agricultural holdings not registered in IACS

Other farms (40 320) from the Statistical Farm Register which data were not linked to the digital data from the AIRBC were also analysed. The addresses of these farms were linked to the Address Register data. For 32 451 farms, this linkage was done successfully and a coordinate of the address was obtained.

Moreover, the farms with an address in largest cities (Vilnius, Kaunas, Klaipėda, Panevėžys, Šiauliai, Alytus) were found. 4 432 farms out of 32 451 were selected. It was decided to try finding another address of these farms, if possible. The databases of the Agricultural Census 2010 and FSS 2013 were used for obtaining the address of the farm centre. For 793 farms the addresses which are not in largest cities were found. For rest of 3 639 farms the addresses in the largest cities were left. In some cases the largest part of the production of these farms can be produced in the cities, because there are functioning farms on the outskirts of the city.

Thus, after the initial analysis, we can predicate that there are 7 869 farms (4.8 per cent of all the farms in the Statistical Farm Register) that have been left without any method for defining geo-coordinates so far. 3 639 farms in largest cities do not belong to these farms. Therefore, other administrative data sources were used for defining addresses of the farm centre of these farms.

First of all, databases of the Agricultural Census 2010 and FSS 2013 were used because during the data collection of these statistical surveys the address of the farm centre (municipality, ward, and locality) were collected and entered into the survey questionnaires. In such a way the addresses of 7 869 farms were taken from 2010 and 2013.

Moreover, after analysis and discussions with experts in Statistics Lithuania, it was decided to try to find the holders of these 7 869 farms in the Real Estate Register. This action was performed and data from the Real Estate Register was obtained. The addresses of all real estate objects registered in the Real Estate Register by the owners of those farms were taken. The structure of data obtained was analyzed as well. The addresses of dwelling houses, farm buildings, yard buildings, barns, etc. were obtained. The addresses were compared to those, obtained from the databases of the Agricultural Census 2010 and FSS 2013. If, according to the Real Estate Register, the farm has several buildings in several municipalities, the main farm building was considered the building in the municipality which was the same as the municipality of the farm centre in the Agricultural Census 2010 and FSS 2013. In such a way the addresses of 7 582 farm buildings (out of 7 869 farms) were obtained. For those farms, which were not found in the Real Estate Register, the farm centre addresses were taken from the databases of the Agricultural Census 2010 and FSS 2013.

Furthermore, a lot of smaller steps were done manually in order to obtain the best result for this project from the data sources which were analysed.

After all steps which were done, the location of farms may be seen in the map as it is shown in Figure 5.

Figure 5. Location of farms
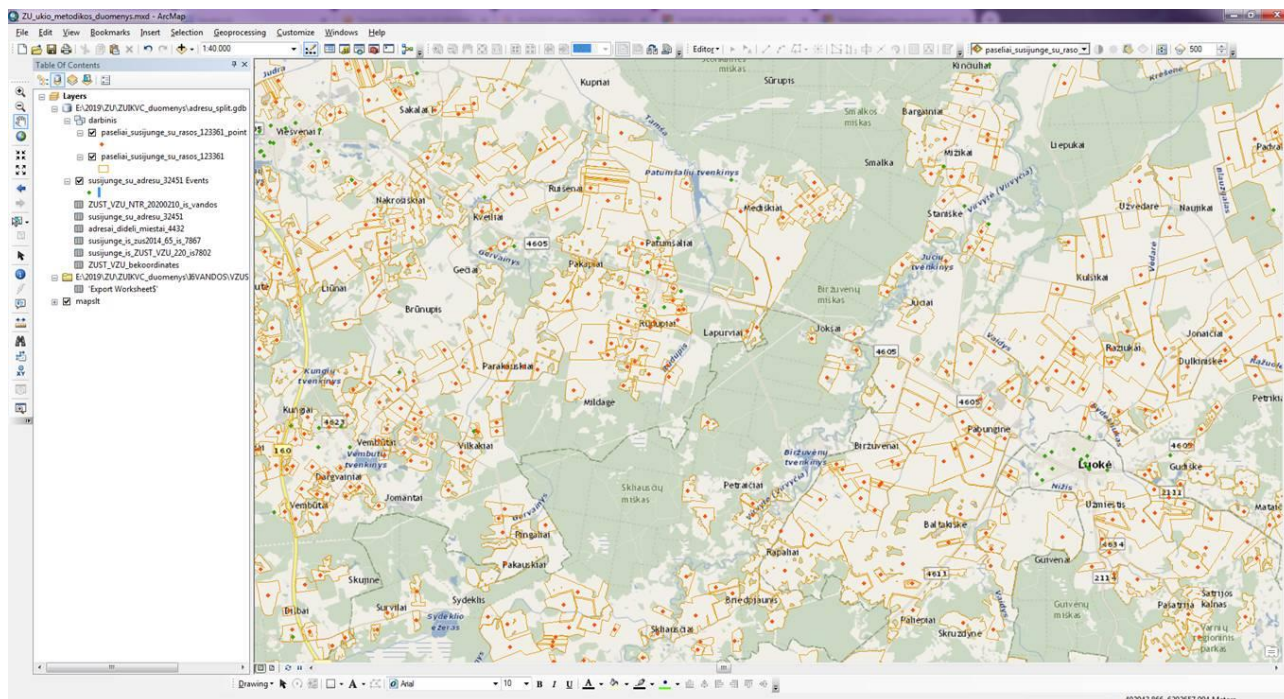
The red points – the farm centres obtained from IACS (the digital data of plots of agricultural land and crop fields declared in IACS in 2018).

The green points – the farm centres obtained from other data sources (Address Register, Real Estate Register, databases of the Agricultural Census 2010 and FSS 2013.

3.2.2. Methodological issues (from the Draft Integrated Farm Statistics Handbook 2020)

According to the Draft Integrated Farm Statistics Handbook (2020), the geographical location of a holding is established according to the following criteria, in order of preference:

1. The location of the main agricultural building.

For livestock farms, it should be where the livestock is located. For crop and mixed farms, it should be the main building of the holding which is usually located close to the agricultural activities.

2. If there is no agricultural building to which a location of the holding could be attributed:

a. The location of the majority of the area of the holding

b. The location of the most important parcel chosen by physical size. The physical size will be calculated on the number of hectares belonging to that parcel which also belongs to the holding located in a particular area.

c. The location of the most important parcel chosen by economic size. The economic size will be calculated on the economic significance of the agricultural production on that parcel. This criterion is more complex due to the variability over the time on the crop sown or to be sown in the selected parcel and should be avoided as IFS is a structural survey and not a production survey. The following

16

order can be used to decide on the significance of the parcel: intensive production on arable land, permanent crops (fruit trees, berries or olive plantations, vineyards) and finally grassland.

3. The location of the farmer's residence if it is not further than 5 km straight from the farm.

4. A combination of the above referred criteria.

5. None of the above referred criteria.

6. The criteria are not applicable (e.g. only one NUTS3 region in the country).

In any case Member States shall provide to Eurostat more detailed information regarding this issue in their quality reports.

3.2.3. Methodology for defining the geographical location of the holdings in the Agricultural Census 2020

During the Agricultural Census 2020, the geographical location of the holdings will be defined according to the Draft Integrated Farm Statistics Handbook (2020). The analysis of the digital data of plots of agricultural land and crop fields declared in IACS in 2018 and data from other sources conducted during this project has shown the possibilities how we can use these data for purposes of the IFS Regulation. This project has shown the way how we can fulfil the requirements of the IFS Regulation regarding geographical location of the farms. This project has shown data sources which can be used and how they can be used.

Step 1. Defining of the main agricultural building of the farm and its geolocation

First of all, the location of the main agricultural building should be obtained. For that purpose, all livestock farms and mixed farms will be selected and checked in the Real Estate Register. All real estate objects registered in the Real Estate Register by the owners of those farms will be obtained. From these real estate objects only farm buildings will he selected.

During the Agricultural Census 2020, in the questionnaire, the farms will have to indicate the municipality in which the farm centre is located. These municipalities will be linked to the data from the Real Estate Register and the farm buildings in these municipalities will be selected from this register. In such a way the main building of each farm will be chosen together with its address.

After that, farms without agricultural building to which location of the holding could be attributed will be defined.

Step 2. Finding the geolocation of farms for which geolocation was not found in the first step

The second step will be finding the geolocation of farms for which geolocation was not found in the first step described above (for livestock farms and mixed farms) and also for remaining farms which have only crop production (crop areas, grassland, permanent crops). The geolocation of farms will be found using the experience gained during this project.

The digital data of plots of agricultural land and crop fields declared in IACS in 2020 will be used. These digital data will be received from the AIRBC. The digital data of all declared in the IACS fields of agricultural areas will be taken. These fields will be linked to number of different owners of farms (agricultural holdings) by using ID code of farm owners. In that way the list of farms which declared their utilised agricultural land in the IACS will be composed and all fields will be linked to these farms.

The farms, geolocation of which was found during the actions described in the first step, will be dropped from this list. In such a way the list of farms which declares their crops in IACS (with all

fields of agricultural area declared by these farms) but geolocation of which was not found during the actions of the first step will be defined. The location of the most important field chosen by physical size will be defined for each farm in this list.

Step 3. Finding the geolocation of the remaining farms

The list of farms, the geolocation of which was not established yet (during step 1 and 2), will be composed, i.e. farms with geolocation will be taken out from the list of farms in the Agricultural Census 2020. For defining the geolocation of these farms, the address of residence will be used. Data will be taken from the Address Register. For those farms, the residence of which will be in the largest cities of Lithuania (if any), the data about the farm centre from the Agricultural Census 2010 and FSS 2023 will be used. However, if these farms will not be found in the databases of Agricultural Census 2010 and FSS 2013, the addresses of their residence will be taken from the Address Register regardless of whether they will be in largest cities or not.

4. SUSTAINABILITY OF THE PROJECT RESULTS

The experience gained during the project and the results achieved will be used for the organisation of statistical surveys and preparation of agricultural statistics in the future.

IFS Regulation states that 'Labour force and Other Gainful Activities' module data collection should be carried out not only in the Agricultural Census 2020, but also in 2023 and 2026. Therefore, the experience gained during this project on using SSIFB data for this module will be applied not only in 2020, but also in 2023 and 2026.

The knowledge gained during the analysis of the data sources and ways how to use these data sources for the establishment of the geographical location of farms will be used in the future. The same way for finding geolocation of farms which was found out during this project will be used in Agricultural Census 2020 and most likely in the data collections for the farm structure statistics in 2023 and 2026. Moreover, the knowledge acquired during the project implementation will be used for the further development of agricultural statistics and for improving the dissemination process.