REPUBLIC OF SLOVENIA
STATISTICAL OFFICE

SURS Litostrojska cesta 54, SI-1000 Ljubljana

Co-funded by
EUROPEAN COMMISSION
EUROSTAT

Modernisation of Agricultural Statistics

Grant Agreement [08441.2017.008-2018.0220] Slovenia

Final Report

Date: 15 May 2020

Version: 0.2

Table of Contents

1 Executive summary - conclusions

The goal of this project (Eurostat grants for 2017; Objective: 08.4 – Provide quality agriculture, fisheries and forestry statistics; Module (DTM): 08.4.11 Agricultural statistics; Title of action: Modernisation of agricultural statistics) is the preparation of the procedures which would contribute to a more efficient implementation of the Integrated Farm Statistics (IFS) in 2020.

The objectives of the project were:

Modification of the Statistical Farm Register (SFR)

- Prepare two masks for viewing individual data that are in the Statistical Farm Register: view with personal data and view without personal data
- Prepare a solution so that sample methodologists can prepare samples without having an insight into personal data
- Define the process of how we can associate the sample (list of FARM_ID) with personal data in the SFR for the needs of creating various address books
- Add new variables to the SFR
- Preparation of procedures for "automatic" update of the SFR
- Review of agricultural holdings in the SFR:
    - "cleaning" duplicate records,
    - searching for links to administrative registers, where the "connection key" does not already exists,
    - searching information of possible new agricultural holdings (horticultural producers).

Converting the administrative farm ID (KMG-MID) to a new statistical FARM_ID

- Preparation of the procedures for translating all administrative farm IDs into statistical farm IDs
- Determine the needs of translating historical data and preparing a timeline for all agricultural surveys containing the administrative farm ID
- Translation of the administrative farm ID into the statistical farm ID across all domains and in the SFR

Harmonization of input data

- Selection of variables important in various agricultural surveys (emphasis on the Farm Structure Survey)
- Determination of the encryption method (standard for all agricultural surveys)
- Implementation in the Farm Structure Surveys (2010, 2013 and 2016)
- Technical harmonization of databases - implementation for the agricultural census 2020; implementation in other surveys after 2020
- Study visit

Preparation for the agricultural census 2020

- Reviewing the new Integrated Farm Survey (IFS) Regulation requirements and defining the data source for each variable in the new IFS Regulation
- Defining variables for which we do not have administrative data, and considering the possibility of model estimation

- An overview of administrative data sources and the identification of possible new ones
- Updating and preparing new agreements (technical protocols) on the transmission of administrative data
- Preparing documentation for each individual administrative source
- Preparing a database with all administrative sources and other auxiliary statistical sources in order to perform the analysis of the quality of the obtained administrative data sources (comparing administrative data with the existing-historical statistical surveys)
- Conducting study visits in countries with experiences in the field of administrative agricultural census

All of the above mentioned objectives were achieved and the procedures were established.

2 Introduction

In statistics it is crucial to develop new ideas to improve the coherence between the subdomains and to ensure the high-quality data for our users. The emerging needs for data can be in some cases fulfilled with administrative data. In that perspective the integration of the administrative data into statistics is vital and needs to be developed on a regular basis, since administrative data are continuously evolving. For that reason, some main activities were planned (task 1 and task 2 of the grant project):

- **Modification of the Statistical Farm Register (SFR) (task 1):**

The Statistical Farm Register in Slovenia was built in 2004 and was modified in 2009.

The Statistical Farm Register (SFR) was established based on data from the Agricultural Census 2000. The main purpose of the SFR is to provide an updated model framework for statistical surveys in the field of agricultural statistics. The SFR includes all agricultural holdings in Slovenia, both family farms and enterprises that meet the statistical definition of an agricultural holding.

The register is designed as an Oracle database, and until August 2009 we were able to view the data through the Oracle application. Starting in January 2010, we can view them through a new application (viewer) made with "Visual Studio". Viewing through the program is only possible at the farm level. If we want to get more farms that meet certain conditions (sample), we can make different prints with programs such as SAS, Access, Oracle, etc.

The SFR contains the address information of the holding and the basic production data we need for sampling. The first sampling frame was created for the survey: "Annual report on livestock production and sown areas in autumn sowing, December 2004". From that point on, sample frames, or address books in the case of the census, are prepared for various statistical surveys in the field of agricultural statistics.

In addition to regular statistical surveys carried out at the Statistical Office of the Republic of Slovenia (hereinafter SURS), SFR is also used in the production of a sample for Farm Accountancy Data Network (FADN) survey, where the Ministry of Agriculture, Forestry and Food (hereinafter MKGP) collects accounting data on income and operations of agricultural holdings.

Table 2.1: List and description of the tables in the Statistical Farm Register (SFR)

TABLE NAME	DESCRIPTION
ADRESARJI	Contains the address data of agricultural holdings and holders
STATISTIKE_KMETIJ	Contains the statistical data (production data) of agricultural holdings
T_ADRESARJI	Contains the address data of agricultural holdings and holders (history data)
T_STATISTIKE_KMETIJ	Contains the statistical data of agricultural holdings (history data)
ADRESARJI_RAZISKAVE	Contains all history data of updating the SFR (administrative source and statistical surveys)
ODGOVORI	Code list for the variable "ODG_ID" - “status of agricultural holding”

PROIZVODNI_TIPI	Code list for the variable "PRO_PRO_ID" – typology codes
RAZISKAVE	Code list for all the surveys and administrative sources that are possible to update the SFR with. It is valid for all variables ending with "…RAZ_ID")
TEZAVNOSTNA_OBMOCJA	Code list for variable "TEZ_OBM" – less favourable areas
TIP_KMETOVANJA	Code list for variable "NACIN_PRIDELAVE" – information on the type of farming (organic, conventional, integrated, farm in transition to organic farming, farm is partially organic farming, farm is partially in transition to organic farming)
VSOTE	The view table contains the SUM for all statistics in the SFR (only for currently active agricultural holdings)

The address table “ADRESARJI” and the statistics table “STATISTIKE_KMETIJ” contain data on agricultural holdings included in the SFR. These two tables reflect the latest state of available data. In the address table “ADRESARJI” there are mostly data on the agricultural holding and the holder and in the statistical table there are variables that show production on the holding (number of livestock, crops, etc.).

The table “ADRESARJI_RAZISKAVE” contains a history of SRKG updates. It tells us with which survey or administrative source an individual agricultural holding was updated. That way we can, for example, track the involvement in statistical surveys (how many times per year).

The table “T_ADRESARJI” contains all history for address data (“T_” meaning history). In the same way the table “T_STATISTIKE_KMETIJ” contains all history data for statistics. Tables are filled automatically with the help of base triggers. Each time a record is changed in the table “ADRESARJI” or in the table “STATISTIKE_KMETIJ”, its new status is also fully recorded in the history. In the tables representing the history of records, one farm may have several records. The number depends on the number of surveys in which the holding participated (as well as the number of updates with administrative registers). Table structures are the same as in the tables with the current state (“ADRESARJI”, “STATISTIKE_KMETIJ”) except that the variable "DATE" is added in the history tables. The value of this variable tells us the date, hour, minute and second of the record’s update. The history records and the current status records are linked to each other by the statistical identifier of agricultural holding (KME_ID, STAT_KME_ID).

Based on new data needs and the tendency for personal data protection (REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), the modification of the Statistical Farm Register was needed in order to provide an appropriate and quality frame for statistical surveys and to provide personal data protection according to the requirements of the applicable legislation. Needed were new variables for stratification purposes (sample design) and redefining the granted accesses to the Statistical Farm Register.

- **Converting the administrative farm ID (KMG-MID) to a statistical farm ID (task 1 / 2):**

The administrative farm ID is considered as a personal identifier and as such needed to be excluded from all databases. For this purpose, the conversion of the administrative farm ID to the statistical ID is needed.

- **Harmonization of input data (task 2):**

The IDs of the variables in previous Farm Structure Surveys (FSS) were designed for each individual survey. The direct comparison through years was for that reason not possible or at least very difficult. The aim is to introduce the harmonized code list for input data of all the main agricultural variables (based on the FSS and the future IFS). In that case the long-time series for trend analysis could be made quickly also on input data. Also, the dissemination of the new tables (long-time series) would be much facilitated. The construction of the questionnaires would become coherent between agricultural domains and cross comparison of the variables collected would be possible.

- **Preparation for the agricultural census 2020 (task 2):**

The aim of the task is to gather (update) all administrative data in the field of agriculture available in Slovenia in order to provide the analysis of the possible administrative census of CORE data in 2020. With the new IFS regulation also model estimation of some variables is a possible way to ensure the data; this way should be examined as well.

3 Project status

3.1 Modification of the Statistical Farm Register (SFR)

3.1.1 OBJECTIVE 1:

Prepare two masks for viewing individual data that are in the Statistical Farm Register on: view with personal data and view without personal data.

PERSONAL DATA PROTECTION ACT

of the Republic of Slovenia (OJ RS, No. 94/07) states:

Personal data - is any data relating to an individual, irrespective of the form in which it is expressed.

Individual - is an identified or identifiable natural person to whom personal data relates; an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity, where the method of identification does not incur large costs or disproportionate effort or require a large amount of time.

Detection of personal data was executed (all are in tables ADRESARJI and T_ADRESARJI).

Table 3.1. List of personal data in the Statistical Farm Register (SFR)

VARIABLE	DESCRIPTION
PRI_IME_NAZ	Name and surname of the holder/manager
TELEFON	Phone number of the holder/manager
TELEFON_2	Second phone number of the holder/manager
MID_HIS_ST	ID number of the holder's/manager's residence
RTE_KME	ID number of the holder's/manager's residence (old version)
RTE_OSB	ID number of the holder's/manager's farm (old version)
MID_OSB	ID number of the holder's/manager's farm
NASLOV	Address of the holder's/manager's residence
SIF_MKG	Administrative farm ID (KMG-MID)
EMAIL	E-mail of the holder/manager
NOVI_MID_HIS_ST	ID number of the holder's/manager's residence (not yet established - future needs)
NOVI_MID_OSB	ID number of the holder's/manager's farm (not yet established - future needs)

Created was a new VIEW table for the table "ADRESARJI" and "T_ADRESARJI" in the Oracle database, since only those two tables contain personal information. The new VIEW tables are called "VIEW_ADRESARJI" and "VIEW_T_ADRESARJI", where all personal data and personal identifiers are removed. In "VIEW" tables also a derived variable was additionally created ("KMG_MID_DA_NE" – if an agricultural holding has an administrative ID ("SIF_MKG") then the variable gets the value "Da" (Yes) and if it does not have it, it gets the value "Ne" (No).

TWO INFORMATION SOURCES WERE CREATED

The first one is “Statistical Farm Register – [STRK@ORA3]” and the second one is “Statistical Farm Register without personal data – [STRK[BOP]@ORA3]”. With first access the personal data can still be accessed (access is assigned only to the SFR subject manager and the SFR developer (access granted also for tables “ADRESARJI” and “T_ADRESARJI”) and the second access without personal data is assigned to all personnel in the statistical office that would need the data for analysis (subject methodologists in other fields of agriculture, sample design team, etc. (access not granted for tables “ADRESARJI” and “T_ADRESARJI”)).

THE NEW VERSION OF THE VIEWER (MASK)

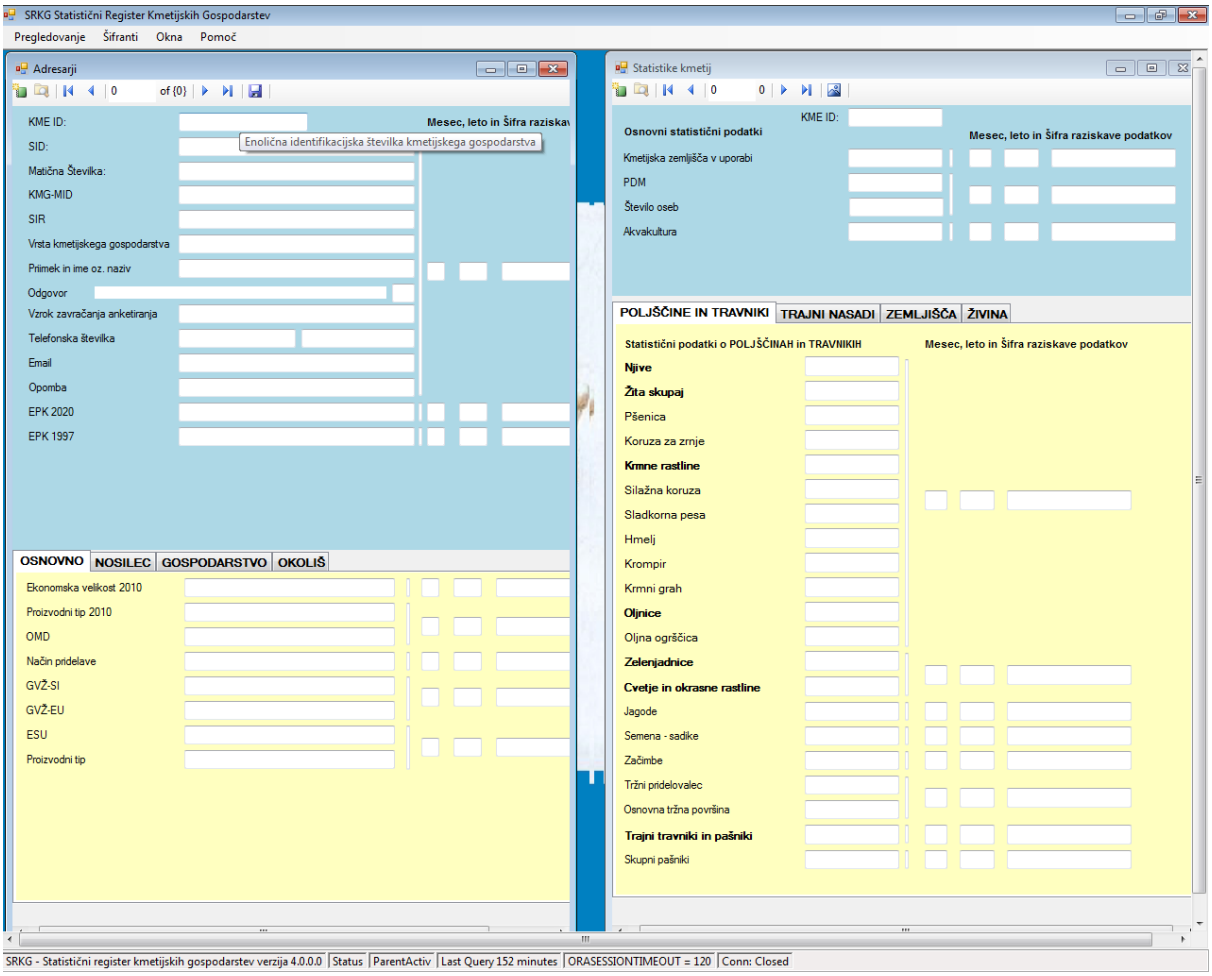
It is made in Visual Studio 2017, with the help of computer language C (see picture 3.1). The previous version of the program did not work anymore, since new variables were added in the SFR database (we had to include them in the viewer). Since there can be two ways of access to information sources, also the two versions of the viewer were created. There is a minor difference between them: one connects to tables with personal data “ADRESARJI” and “T_ADRESARJI” and the other to tables without personal data “VIEW_ADRESARJI” and “VIEW_T_ADRESARJI”, otherwise the programs are identical and work in the same way:

- The view of the farm is possible on an individual level.
- There are many search options when looking for a group of specific farms ”>, <” – larger, smaller.
- It is also possible to search farms with function ”contains” %text%.
- When using the conditions, the “viewer” returns the first found farm on the screen and writes the number of agricultural holdings fitting the conditions (all can be examined individually).
- It allows to check the “code lists” and also the “totals” of statistical data.
- Each variable has an explanation (description) which is visible when putting the mouse pointer on the variable name.

ACCESS TO PERSONAL DATA IS RECORDED

When connecting to information source “Statistical Farm Register – [STRK@ORA3]”- with personal data, access is recorded. This was developed in .NET environment. The application uses PL / SQL procedures. These procedures record each access to the data and record each query/condition that has been executed in a dedicated table. In the same way, access to data through program (SAS) is recorded.

Picture 3.1. The new version of the viewer (mask)



Shown are the two sheets of the viewer; on the first sheet we can see all the data for the farm that are in the table "ADRESARJI" or "VIEW_ADRESARJI" and on the second sheet we can see all the statistics for the farm that we would select.

3.1.2 OBJECTIVE 2:

Prepare a solution so that sample methodologists can prepare samples without having an insight into personal data

The solution in objective 1 (see "objective 1" of "3.1 Modification of the Statistical Farm Register (SFR)") was used also for objective 2. When a sample methodologist is developing a new sample for agricultural surveys, the access to information source without personal data – [STRK[BOP]@ORA3] is granted. That way the sample is made without accessing the personal data but having access to all statistics needed to create a sample.

3.1.3 OBJECTIVE 3:

Define the process of how we can associate the sample (list of FARM_IDs) with personal data in the Statistical Farm Register for the needs of creating various address books

The solution in objective 1 (see "objective 1" of "3.1 Modification of the Statistical Farm Register (SFR)") was used also for objective 3. When a sample methodologist develops a

new sample for agricultural surveys (see “objective 2” of “3.1 Modification of the Statistical Farm Register (SFR)), the list of farms (FARM_ID = KME_ID) is sent to the department working on addresses. The access to the information source with personal data – [STRK@ORA3] is granted for a short period of time and the list of farms is linked (through KME_ID) with personal data within the database. The variables needed in the address book are predefined. The full address book is then put away to a secured folder and used for collecting the data from agricultural holdings.

3.1.4 OBJECTIVE 4:

Add new variables to the Statistical Farm Register

While examining the list of variables in the SFR, we encountered some variables that were added when the SFR was built (in 2004), but were not used - those were removed (SOC_SOC_ID, SOC_SPR_M, SOC_SPR_L, SOC_SPR_RAZ_ID). Removed was also the variable NAMAKANJE_Z, which indicated if an agricultural holding irrigates the land in a greenhouse (this variable actually indicated if an agricultural holding has a greenhouse or not). This variable was removed, since it was supplemented with a new variable (ZASC_PROSTOR), which indicates the area under greenhouse.

From table “ADRESARJI” 4 variables were deleted and 8 variables were added.

From table “STATISTIKE_KMETIJ” 1 variable was deleted and 58 were added (out of those 16 were actual statistics and other 42 were so-called supplementing variables (variables that indicate the month, year and the code of the survey where the variable was changed)).

Besides adding new variables also one CODE list was modified (valid for variable TIP_KMETOVANJA). Besides organic, conventional, integrated, farm in transition to organic farming, added were also codes for farm is partially organic farming, and farm is partially in transition to organic farming.

Table 3.2. List of variables added into the Statistical Farm Register (SFR)

NAME OF THE VARIABLE	**LOCATION OF THE TABLE IN SFR**	**DESCRIPTION OF THE VARIABLE**
TIP_KMETIJE	ADRESARJI	Type of agricultural holding as described in the administrative register of agricultural holdings, established by the Ministry of Agriculture, Forestry and Food (MKGP) (legal entity, family farm, agrarian community, grazing community)
SID	ADRESARJI	Statistical personal identification number
SIR	ADRESARJI	Statistical legal entity identification number
OPOMBA	ADRESARJI	Remarks made for a specific agricultural holding
NAC_ZAVRACANJA	ADRESARJI	Reason for non-response in statistical surveys (rejecting to give an answer, non-availability to provide an answer (death in family, illness, age, not understanding the language), non-response when making a call to a family farm, no telephone number for family farm)
EMAIL	ADRESARJI	EMAIL address of the holder/manager of the farm

NOVI_MID_HIS_ST	ADRESARJI	Possible new house ID number of the holder/manager (not yet established but in process)
NOVI_MID_OSB	ADRESARJI	Possible new house ID number of the farm (not yet established but in process)
GOBE	STATISTIKE_KMETIJ	Area of mushrooms
GOBE_SPR_M	STATISTIKE_KMETIJ	Month of change in the variable GOBE
GOBE_SPR_L	STATISTIKE_KMETIJ	Year of change in the variable GOBE
GOBE_RAZ_ID	STATISTIKE_KMETIJ	The code of the survey where the variable GOBE was changed
ZASC_PROSTOR	STATISTIKE_KMETIJ	Area under greenhouses
ZASC_SPR_M	STATISTIKE_KMETIJ	Month of change in the variable ZASC_PROSTOR
ZASC_SPR_L	STATISTIKE_KMETIJ	Year of change in the variable ZASC_PROSTOR
ZASC_RAZ_ID	STATISTIKE_KMETIJ	The code of the survey where the variable ZASC_PROSTOR was changed
DREVESNICE	STATISTIKE_KMETIJ	The total area of nurseries
DREV_SPR_M	STATISTIKE_KMETIJ	Month of change in the variable DREVESNICE
DREV_SPR_L	STATISTIKE_KMETIJ	Year of change in the variable DREVESNICE
DREV_RAZ_ID	STATISTIKE_KMETIJ	The code of the survey where the variable DREVESNICE was changed
ZACIMBE	STATISTIKE_KMETIJ	The area of spices, condiments and medicinal plants
ZAC_SPR_M	STATISTIKE_KMETIJ	Month of change in the variable ZACIMBE
ZAC_SPR_L	STATISTIKE_KMETIJ	Year of change in the variable ZACIMBE
ZAC_RAZ_ID	STATISTIKE_KMETIJ	The code of the survey where the variable ZACIMBE was changed
EKST_SAD	STATISTIKE_KMETIJ	The area of extensive orchards
JAGODE	STATISTIKE_KMETIJ	The area of strawberries
JAGODE_SPR_M	STATISTIKE_KMETIJ	Month of change in the variable JAGODE
JAGODE_SPR_L	STATISTIKE_KMETIJ	Year of change in the variable JAGODE

JAGODE_RAZ_ID	STATISTIKE_KMETIJ	The code of the survey where the variable JAGODE was changed
SEMENA_SADIKE	STATISTIKE_KMETIJ	The area of seeds and seedlings
SEM_SAD_SPR_M	STATISTIKE_KMETIJ	Month of change in the variable SEMENA_SADIKE
SEM_SAD_SPR_L	STATISTIKE_KMETIJ	Year of change in the variable SEMENA_SADIKE
SEM_SAD_RAZ_ID	STATISTIKE_KMETIJ	The code of the survey where the variable SEMENA_SADIKE was changed
SKUPNI_PASNIK	STATISTIKE_KMETIJ	The area of common grassland
SKUP_PAS_SPR_M	STATISTIKE_KMETIJ	Month of change in the variable SKUPNI_PASNIK
SKUP_PAS_SPR_L	STATISTIKE_KMETIJ	Year of change in the variable SKUPNI_PASNIK
SKUP_PAS_RAZ_ID	STATISTIKE_KMETIJ	The code of the survey where the variable SKUPNI_PASNIK was changed
KZU_EKO	STATISTIKE_KMETIJ	Utilized agricultural area of organic farming
KZU_PREUSMERITEV	STATISTIKE_KMETIJ	Utilized agricultural area in conversion to organic farming
EKO_SPR_M	STATISTIKE_KMETIJ	Month of change in the variable KZU_EKO, KZU_PREUSMERITEV
EKO_SPR_L	STATISTIKE_KMETIJ	Year of change in the variable KZU_EKO, KZU_PREUSMERITEV
EKO_RAZ_ID	STATISTIKE_KMETIJ	The code of the survey where the variable KZU_EKO, KZU_PREUSMERITEV was changed
GVZ_EU	STATISTIKE_KMETIJ	Livestock size unit, calculated based on EU livestock unit coefficients
POLZI	STATISTIKE_KMETIJ	The amount of snails (in kg)
POLZI_SPR_M	STATISTIKE_KMETIJ	Month of change in the variable POLZI
POLZI_SPR_L	STATISTIKE_KMETIJ	Year of change in the variable POLZI
POLZI_RAZ_ID	STATISTIKE_KMETIJ	The code of the survey where the variable POLZI was changed
KR_MLEKO	STATISTIKE_KMETIJ	Annual amount of purchased (sold from farmers) cow’s milk
KR_MLEKO_SPR_M	STATISTIKE_KMETIJ	Month of change in the variable KR_MLEKO
KR_MLEKO_SPR_L	STATISTIKE_KMETIJ	Year of change in the variable KR_MLEKO
KR_MLEKO_RAZ_ID	STATISTIKE_KMETIJ	The code of the survey where the variable KR_MLEKO was changed

PDM	STATISTIKE_KMETIJ	Annual working unit (AWU)
ST_OSEB	STATISTIKE_KMETIJ	Number of persons regularly working on the farm
PDM_SPR_M	STATISTIKE_KMETIJ	Month of change in the variable PDM, ST_OSEB
PDM_SPR_L	STATISTIKE_KMETIJ	Year of change in the variable PDM, ST_OSEB
PDM_RAZ_ID	STATISTIKE_KMETIJ	The code of the survey where the variable PDM, ST_OSEB was changed
AKVAKULTURA	STATISTIKE_KMETIJ	Does the agricultural holding deal with aquaculture
AKVA_SPR_M	STATISTIKE_KMETIJ	Month of change in the variable AKVAKULTURA
AKVA_SPR_L	STATISTIKE_KMETIJ	Year of change in the variable AKVAKULTURA
AKVA_RAZ_ID	STATISTIKE_KMETIJ	The code of the survey where the variable AKVAKULTURA was changed
GOZD_SPR_M	STATISTIKE_KMETIJ	Month of change in the variable GOZD
GOZD_SPR_L	STATISTIKE_KMETIJ	Year of change in the variable GOZD
GOZD_RAZ_ID	STATISTIKE_KMETIJ	The code of the survey where the variable GOZD was changed
KZU_SPR_M	STATISTIKE_KMETIJ	Month of change in the variable KME_ZEM_UPO
KZU_SPR_L	STATISTIKE_KMETIJ	Year of change in the variable KME_ZEM_UPO
KZU_RAZ_ID	STATISTIKE_KMETIJ	The code of the survey where the variable KME_ZEM_UPO was changed

Since some of the added variables were directly correlated to the Farm Structure Survey (FFS) form previous years, we updated the Statistical Farm Register directly from those databases; for 2010, 2013 and 2016 all agricultural holdings being eligible in the mentioned surveys got an update for variables: GVZ_EU, GOBE, ZASC_PROSTOR, DREVESNICE, ZACIMBE, JAGODE, SEMENA_SADIKE, SKUPNI_PASNIK, PDM, ST_OSEB, AKVAKULTURA, INT_TRAVNIK, EKST_TRAVNIK, KRA_MOLZNICE, KRA_DOJILJE. With those, the so-called supplementing variables (variables that indicate the month, year and the code of the survey where the variable was changed) were also updated.

3.1.5 OBJECTIVE 5:

Preparation of procedures for "automatic" update of the Statistical Farm Register

"Automatic" update of the Statistical Farm Register is made in Oracle (SQL environment).

The Statistical Farm Register is regularly updated two times per year (in February/March and in September/October) and most of the update is derived directly from the administrative source (with direct access), so it is sensible to use “automatic” update. When we speak about “automatic” update, it is considered that most of the tasks are triggered automatically with a “push of a button”, but nevertheless manual examinations (interim outputs) are still needed in order to detect possible changes in the administrative register (administrative registers are constantly evolving) or some mistakes when the data are inserted into the register.

The Statistical Office of the Republic of Slovenia has a granted direct access to the administrative source. It is a direct access to their database with username and password. The access is limited and the users (subject methodologist and IT staff) had to sign a special document specifying proper handling with personal data.

For updating, some administrative registers are used with direct access:

- Register of Agricultural Holdings (RAH), managed by the Ministry of Agriculture, Forestry and Food (MKGP) (integrated are data on holder/manager, orchards, strawberries, vineyard, olive trees, area under greenhouse)
- Register of Beehives, managed by the Agency for Agricultural Markets and Rural Development (data on beehives)
- Register of Cattle, managed by the Agency for Agricultural Markets and Rural Development (data on cattle)
- Register of Sheep and Goats, managed by the Agency for Agricultural Markets and Rural Development (data on sheep and goats)
- Register of Equidae, managed by the Agency for Agricultural Markets and Rural Development (data on equidae)
- Register of Pigs, managed by the Agency for Agricultural Markets and Rural Development (data on pigs)
- Register of Territorial Units, managed by the Surveying and Mapping Authority of the Republic of Slovenia (data on location)
- Central Population Register, managed by the Ministry of the Interior (data on holder/manager and their family)
- Statistical Business Register (internal statistical source; Framework Regulation Integrating Business Statistics (FRIBS))

For updating, also Integrated Administration and Control System (IACS) data are used; data when farmers apply for subsidies (data on land parcel area by different crop and number of livestock on agricultural holding). These data are received by the Statistical Office as a pre-prepared dataset through the Secured File Transfer Protocol (SFTP) once a year and are used for updating the Statistical Farm Register (SFR) in September/October. For updating in February/March this dataset is not used since the data are not available at that time.

Each individual register has its own specifics. They were examined and the SQL codes were constructed so that from each register valid data were derived (precise instructions from subject methodologist to IT were created and are kept in the Statistical Office as a working document).

The first step is to link all administrative sources with the Statistical Farm Register (SFR) into one master table (the basis is the agricultural holdings in the administrative register of agricultural holdings, since this is the population of interest). Linking is not problematic, since

all of the above mentioned registers (datasets) have a least one identifier, which is uniquely assigned to location, person, agricultural holding, business. The critical date when the extraction is made is the 1st of the month when the update is being executed (1 February or March and 1 September or October).

In the second step interim outputs are developed for manual examination by the subject methodologist. Several files are created based on pre-prepared conditions:

1. We examine the agricultural holdings where in the administrative Register of Agricultural Holdings (RAH) the variable VRSTA is not equal to “Kme” or “KMG”. Meaning the output of the agricultural holdings that are not family farms or agricultural enterprises. For now those are only “Common land agricultural holdings”. Those are not updated in the Statistical Farm Register (SFR) since the area of common land is distributed to agricultural holdings which have animals grazing the common pastures (when the survey is conducted).
2. In the administrative Register of Agricultural Holdings (RAH) there are still some fiction location IDs for holder/manager location or for location of the farm. In most cases the holder/manager lives or has a farm in a different country. Here can also be farms that have old ID numbers that are not valid (for those we can manually assign location (village, municipality, etc.)).
3. We determined some agricultural holdings - enterprises which have some area of agricultural land or animals, but their output is not for agricultural purposes. Listed are schools, hunting societies, religious community, zoo, retirement homes, psychiatric hospitals, etc. The area or the number of animals are negligible; for now we treat them as not-eligible agricultural holdings. For some this might change with the agricultural census 2020.
4. We examine all agricultural holdings that are currently in the Statistical Farm Register (SFR) stated as non-eligible (under the threshold or not a farm anymore), but from the administrative data we see that there is production above the threshold. Those are first linked to agricultural holdings that do not have an administrative ID in the Statistical Farm Register (SFR) (meaning SIF_MKG is missing; see “objective 6” of “3.1 Modification of the Statistical Farm Register (SFR)”). The linkage is made through location IDs and name/surname. That way we can attribute administrative IDs (SIF_MKG) to existing agricultural holdings in the SFR. All others are then considered as new “eligible farms”.
5. We examine the ID number of the enterprise in the Statistical Farm Register (SFR) where it is not the same as in the administrative Register of Agricultural Holdings (RAH) (derived from the business register). The first reason for this is that some enterprises have subunits which deal with agriculture but the ID of the main unit is written in RAH (which has a different line of work). We manually change some business IDs in the SFR, since we want to send the mail to enterprises (subunits) where agriculture is managed. The second reason can occur when one agricultural holding changes the status from a family farm to an enterprise.
6. We examine agricultural holdings that stated in the administrative Register of Agricultural Holdings that they are not eligible any more. In those cases we change the status in the Statistical Farm Register (SFR) from eligible to non-eligible.
7. We examine agricultural holdings where the holder/manager is not alive. From administrative source “Central population register (CRP)” all citizens that are on a critical date (just before the update of the SFR) alive, are taken. Those who are then linked with SFR (holders/mangers) are considered alive. Those who do not link, are considered as dead holders/managers; family farms get a special response number in variable ODG_ID (from eligible farm to eligible farm where the holder is dead). When we make address books for statistical surveys, we address the mail only to the surname (we delete the name) in order not to cause any additional inconvenience in

those difficult situations. During the survey we also ask the question if a certain person is still a holder/manager of the farm. The agricultural house holding has a possibility to answer “No” and one of the reasons can be that the farmer/holder is dead (we ask for the new holder/manager’s name). This new information is at the end of the statistical survey updated in Statistical Farm Register (SFR).

8. A final list of agricultural holdings considered being updated from administrative data is extracted for examination.

Table 3.3.Variables updated with “automatic” update in the Statistical Farm Register (SFR); table ADRESARJI

VARIABLE	**SOURCE**	**DESCRIPTION**
SIF_GOSP_MAT_ST	Statistical Business Register	Business ID number
ODG_ID	Register of Agricultural Holdings (RAH) and Central Population Register	Response ID (eligible farm, non-eligible farm, eligible farm, where the holder/manager is dead)
PRI_IME_NAZ	Register of Agricultural Holdings	Name and surname of the holder/manager
PRI_IME_SPR_M	Register of Agricultural Holdings	Month of change in the variable PRI_IME_NAZ
PRI_IME_SPR_L	Register of Agricultural Holdings	Year of change in the variable PRI_IME_NAZ
PRI_IME_RAZ_ID	Register of Agricultural Holdings	The code of the survey where the variable PRI_IME_NAZ was changed
TELEFON	Register of Agricultural Holdings (RAH)	Phone number of the holder/manager
MID_HIS_ST	Register of Territorial Units	ID number of the holder’s/manager’s residence
RTE_MID_KME_SPR_M	Register of Territorial Units	Month of change in the variable MID_HIS_ST
RTE_MID_KME_SPR_L	Register of Territorial Units	Year of change in the variable MID_HIS_ST
RTE_MID_KME_RAZ_ID	Register of Territorial Units	The code of the survey where the variable MID_HIS_ST was changed
MID_OSB	Register of Territorial Units	ID number of the holder’s/manager’s farm
RTE_MID_OSB_SPR_M	Register of Territorial Units	Month of change in the variable MID_OSB
RTE_MID_OSB_SPR_L	Register of Territorial Units	Year of change in the variable MID_OSB
RTE_MID_OSB_RAZ_ID	Register of Territorial Units	The code of the survey where the variable MID_OSB was changed
NASELJE_ID	Register of Territorial Units	ID of village for the location of the holder’s/manager’s residence

NASELJE_IME	Register of Territorial Units	Name of village for the location of the holder's/manager's residence
OBCINA_ID	Register of Territorial Units	ID of municipality for the location of the holder's/manager's residence
OBCINA_IME	Register of Territorial Units	Name of municipality for the location of the holder's/manager's residence
REGIJA_ID	Register of Territorial Units	ID of NUTS 3 region for the location of the holder's/manager's residence
REGIJA_IME	Register of Territorial Units	Name of NUTS 3 region for the location of the holder's/manager's residence
GEO_UMESTITEV_M	Register of Territorial Units	Month of change in the variables NASELJE_ID, NASELJE_IME, OBCINA_ID, OBCINA_IME, REGIJA_ID, REGIJA_IME
GEO_UMESTITEV_L	Register of Territorial Units	Year of change in the variables NASELJE_ID, NASELJE_IME, OBCINA_ID, OBCINA_IME, REGIJA_ID, REGIJA_IME
GEO_UMESTITEV_RAZ_ID	Register of Territorial Units	The code of the survey where the variables: NASELJE_ID, NASELJE_IME, OBCINA_ID, OBCINA_IME, REGIJA_ID, REGIJA_IME, were changed
NASELJE_ID_SDZ	Register of Territorial Units	ID of village for the location of agricultural holding
NASELJE_IME_SDZ	Register of Territorial Units	Name of village for the location of agricultural holding
OBCINA_ID_SDZ	Register of Territorial Units	ID of municipality for the location of agricultural holding
OBCINA_IME_SDZ	Register of Territorial Units	Name of municipality for the location of agricultural holding
REGIJA_ID_SDZ	Register of Territorial Units	ID of NUTS 3 region for the location of agricultural holding
REGIJA_IME_SDZ	Register of Territorial Units	Name of NUTS 3 region for the location of agricultural holding
GEO_UME_SDZ_M	Register of Territorial Units	Month of change in the variables NASELJE_ID_SDZ, NASELJE_IME_SDZ, OBCINA_ID_SDZ, OBCINA_IME_SDZ, REGIJA_ID_SDZ, REGIJA_IME_SDZ
GEO_UME_SDZ_L	Register of Territorial Units	Year of change in the variables NASELJE_ID_SDZ, NASELJE_IME_SDZ, OBCINA_ID_SDZ, OBCINA_IME_SDZ, REGIJA_ID_SDZ, REGIJA_IME_SDZ
GEO_UME_SDZ_RAZ_ID	Register of Territorial Units	The code of the survey where the variables NASELJE_ID_SDZ, NASELJE_IME_SDZ,

		OBCINA_ID_SDZ, OBCINA_IME_SDZ, REGIJA_ID_SDZ, REGIJA_IME_SDZ were changed
NASLOV	Register of Territorial Units	Address of the holder's/manager's residence
POSTNA_ST	Register of Territorial Units	Post number of the holder's/manager's residence
POSTA	Register of Territorial Units	Name of the post of the holder's/manager's residence
SIF_MKG	Register of Agricultural Holdings (RAH)	Administrative farm ID (KMG-MID)
TIP_KMETIJE	Register of Agricultural Holdings (RAH)	Type of agricultural holding as described in the Administrative Register of Agricultural Holdings, established by the Ministry of Agriculture Forestry and Food (MKGP) (legal entity, family farm, agrarian community, grazing community)
SID	Central Population Register	Statistical personal identification number (translated from personal identification number)
SIR	Statistical Business Register	Statistical legal entity identification number (translated from personal identification number)
EMAIL	Register of Agricultural Holdings (RAH)	E-mail of holder/manager

Table 3.4. Variables updated with "automatic" update in the Statistical Farm Register (SFR), table STATISTIKE_KMETIJ.

VARIABLE	**SOURCE**	**DESCRIPTION**
TRZ_PRID_RAZ_ID	Register of Agricultural Holdings (RAH)	The code of the survey where the variable TRZ_PRID was changed
TRZ_PRID_SPR_M	Register of Agricultural Holdings (RAH)	Month of change in the variable TRZ_PRID
TRZ_PRID_SPR_L	Register of Agricultural Holdings (RAH)	Year of change in the variable TRZ_PRID
TRZ_PRID	Register of Agricultural Holdings (RAH)	Variable is updated only when the area of strawberries in the register is larger than 0.1 hectare. For those agricultural holdings we are certain that they are market producers
ZIT_SK	Integrated Administration and Control System (IACS)	Area of cereals for the production of grain
KRM_RAS	Integrated Administration and Control System (IACS)	Area of forage plants
PSENICA	Integrated Administration and Control System (IACS)	Area of wheat

SIL_KOR	Integrated Administration and Control System (IACS)	Area of green maize
SLA_PES	Integrated Administration and Control System (IACS)	Area of sugar beet
KOR_ZRN	Integrated Administration and Control System (IACS)	Area of grain maize
HMELJ	Register of Agricultural Holdings (RAH)	Area of hops
KROMPIR	Integrated Administration and Control System (IACS)	Area of potato
OLJNICE	Integrated Administration and Control System (IACS)	Area of oilseeds
NJIVE	Integrated Administration and Control System (IACS)	Area of arable land
POLJ_SPR_M	Integrated Administration and Control System (IACS)	Month of change in the variables ZIT_SK, KRM_RAS, PSENICA, SIL_KOR, SLA_PES, KOR_ZRN, HMELJ, KROMPIR, OLJNICE, NJIVE
POLJ_SPR_L	Integrated Administration and Control System (IACS)	Year of change in the variables ZIT_SK, KRM_RAS, PSENICA, SIL_KOR, SLA_PES, KOR_ZRN, HMELJ, KROMPIR, OLJNICE, NJIVE
POLJ_RAZ_ID	Integrated Administration and Control System (IACS)	The code of the survey where the variables ZIT_SK, KRM_RAS, PSENICA, SIL_KOR, SLA_PES, KOR_ZRN, HMELJ, KROMPIR, OLJNICE, NJIVE were changed
VIN_SK	Register of Agricultural Holdings (RAH)	Area on vineyard
VIN_SPR_M	Register of Agricultural Holdings (RAH)	Month of change in the variable VIN_SK
VIN_SPR_L	Register of Agricultural Holdings (RAH)	Year of change in the variable VIN_SK
VIN_RAZ_ID	Register of Agricultural Holdings (RAH)	The code of the survey where the variable VIN_SK was changed
INT_SAD	Register of Agricultural Holdings (RAH)	Area of intensive orchards
INT_SAD_SPR_M	Register of Agricultural Holdings (RAH)	Month of change in the variable INT_SAD
INT_SAD_SPR_L	Register of Agricultural Holdings (RAH)	Year of change in the variable INT_SAD
INT_SAD_RAZ_ID	Register of Agricultural Holdings (RAH)	The code of the survey where the variable INT_SAD was changed
TRA_PAS	Integrated Administration	Area of permanent grassland

	and Control System (IACS)	
TRA_PAS_SPR_M	Integrated Administration and Control System (IACS)	Month of change in the variable TRA_PAS
TRA_PAS_SPR_L	Integrated Administration and Control System (IACS)	Year of change in the variable TRA_PAS
TRA_PAS_RAZ_ID	Integrated Administration and Control System (IACS)	The code of the survey where the variable TRA_PAS was changed
ST_DREVES	Register of Agricultural Holdings (RAH)	Number of trees in extensive orchards
ST_DREV_SPR_M	Register of Agricultural Holdings (RAH)	Month of change in the variable TRA_PAS
ST_DREV_SPR_L	Register of Agricultural Holdings (RAH)	Year of change in the variable TRA_PAS
ST_DREV_RAZ_ID	Register of Agricultural Holdings (RAH)	The code of the survey where the variable TRA_PAS was changed
ZELENJ	Integrated Administration and Control System (IACS)	Area of fresh vegetables
CVE_OKR_RAST	Integrated Administration and Control System (IACS)	Area of flowers and ornamental plants (excluding nurseries)
ZEL_CVE_OKR_SPR_M	Integrated Administration and Control System (IACS)	Month of change in the variables ZELENJ, CVE_OKR_RAST
ZEL_CVE_OKR_RAZ_ID	Integrated Administration and Control System (IACS)	Year of change in the variables ZELENJ, CVE_OKR_RAST
ZEL_CVE_OKR_SPR_L	Integrated Administration and Control System (IACS)	The code of the survey where the variables ZELENJ, CVE_OKR_RAST was changed
KME_ZEM_UPO	Integrated Administration and Control System (IACS)	Utilized agricultural area
KON_SK	Register of Equidae	Number of equidae
KON_SPR_M	Register of Equidae	Month of change in the variable KON_SK
KON_SPR_L	Register of Equidae	Year of change in the variable KON_SK
KON_RAZ_ID	Register of Equidae	The code of the survey where the variable KON_SK was changed
GOV_SK	Register of Cattle	Number of cattle – total
KRA_SK	Register of Cattle	Number of cows
GOV_SPR_M	Register of Cattle	Month of change in the variables GOV_SK, KRA_SK
GOV_SPR_L	Register of Cattle	Year of change in the variables GOV_SK, KRA_SK
GOV_RAZ_ID	Register of Cattle	The code of the survey where the variables GOV_SK, KRA_SK were changed
PRA_SK	Register of Pigs	Number of pigs – total

PLE_SVI_MLA	Register of Pigs	Number of breeding sows
PRA_PIT	Register of Pigs	Number of pigs for fattening
PRA_SPR_M	Register of Pigs	Month of change in the variables PRA_SK, PLE_SVI_MLA, PRA_PIT
PRA_SPR_L	Register of Pigs	Year of change in the variables PRA_SK, PLE_SVI_MLA, PRA_PIT
PRA_RAZ_ID	Register of Pigs	The code of the survey where the variables PRA_SK, PLE_SVI_MLA, PRA_PIT were changed
PER_SK	Integrated Administration and Control System (IACS)	Number of poultry – total
PIT_PIS	Integrated Administration and Control System (IACS)	Number of broilers
KOK_NES	Integrated Administration and Control System (IACS)	Number of laying hens
INT_REJ_PER	Integrated Administration and Control System (IACS)	Indicator if the agricultural holding is an intensive breeder of poultry (all that have at least 500 poultry)
INT_REJ_PER_M	Integrated Administration and Control System (IACS)	Month of change in the variable INT_REJ_PER
INT_REJ_PER_L	Integrated Administration and Control System (IACS)	Year of change in the variable INT_REJ_PER
INT_REJ_PER_RAZ_ID	Integrated Administration and Control System (IACS)	The code of the survey where the variable INT_REJ_PER was changed
PER_SPR_M	Integrated Administration and Control System (IACS)	Month of change in the variables PER_SK, PIT_PIS, KOK_NES
PER_SPR_L	Integrated Administration and Control System (IACS)	Year of change in the variables PER_SK, PIT_PIS, KOK_NES
PER_RAZ_ID	Integrated Administration and Control System (IACS)	The code of the survey where the variables PER_SK, PIT_PIS, KOK_NES were changed
OVC_SK	Register of Sheep and Goats	Number of sheep – total
PLE_OVC	Register of Sheep and Goats	Number of sheep - breeding females
OVC_SPR_M	Register of Sheep and Goats	Month of change in the variable TRA_PAS
OVC_SPR_L	Register of Sheep and Goats	Year of change in the variable TRA_PAS
OVC_RAZ_ID	Register of Sheep and Goats	The code of the survey where the variable TRA_PAS was changed
KOZ_SK	Register of Sheep and Goats	Number of goats - total
PLE_KOZ	Register of Sheep and Goats	Number of goats - breeding females
KOZ_SPR_M	Register of Sheep and	Month of change in the variables

	Goats	OVC_SK, PLE_OVC, KOZ_SK, PLE_KOZ.
KOZ_SPR_L	Register of Sheep and Goats	Year of change in the variables OVC_SK, PLE_OVC, KOZ_SK, PLE_KOZ
KOZ_RAZ_ID	Register of Sheep and Goats	The code of the survey where the variables OVC_SK, PLE_OVC, KOZ_SK, PLE_KOZ were changed
NOJI	Integrated Administration and Control System (IACS)	Number of ostriches
NOJI_SPR_M	Integrated Administration and Control System (IACS)	Month of change in the variable NOJI
NOJI_SPR_L	Integrated Administration and Control System (IACS)	Year of change in the variable NOJI
NOJI_RAZ_ID	Integrated Administration and Control System (IACS)	The code of the survey where the variable NOJI was changed
PREPELJICE	Integrated Administration and Control System (IACS)	Number of quail
PREPELJICE_SPR_M	Integrated Administration and Control System (IACS)	Month of change in the variable PREPELJICE
PREPELJICE_SPR_L	Integrated Administration and Control System (IACS)	Year of change in the variable PREPELJICE
PREPELJICE_RAZ_ID	Integrated Administration and Control System (IACS)	The code of the survey where the variable PREPELJICE was changed
KUNCI	Integrated Administration and Control System (IACS)	Number of rabbits – total
KUNCI_SPR_M	Integrated Administration and Control System (IACS)	Month of change in the variable KUNCI
KUNCI_SPR_L	Integrated Administration and Control System (IACS)	Year of change in the variable KUNCI
KUNCI_RAZ_ID	Integrated Administration and Control System (IACS)	The code of the survey where the variable KUNCI was changed
JELENJAD	Integrated Administration and Control System (IACS)	Number of deer
JELENJAD_SPR_M	Integrated Administration and Control System (IACS)	Month of change in the variable JELENJAD
JELENJAD_SPR_L	Integrated Administration and Control System (IACS)	Year of change in the variable JELENJAD
JELENJAD_RAZ_ID	Integrated Administration	The code of the survey where the

	and Control System (IACS)	variable JELENJAD was changed
CEBELJE_DRUZINE	Register of Beehives	Number of beehives
CEBELE_SPR_M	Register of Beehives	Month of change in the variable CEBELJE_DRUZINE
CEBELE_SPR_L	Register of Beehives	Year of change in the variable CEBELJE_DRUZINE
CEBELE_RAZ_ID	Register of Beehives	The code of the survey where the variable CEBELJE_DRUZINE was changed
OLJCNIKI_SKUPAJ	Register of Agricultural Holdings (RAH)	Are of olive trees
OLJNA_OGRSCICA	Integrated Administration and Control System (IACS)	Area of rape and turnip rape seeds
KRMNI_GRAH	Integrated Administration and Control System (IACS)	Area of field peas
ZASC_PROSTOR	Register of Agricultural Holdings (RAH)	Area under greenhouses
ZASC_SPR_M	Register of Agricultural Holdings (RAH)	Month of change in the variable ZASC_PROSTOR
ZASC_SPR_L	Register of Agricultural Holdings (RAH)	Year of change in the variable ZASC_PROSTOR
ZASC_RAZ_ID	Register of Agricultural Holdings (RAH)	The code of the survey where the variable ZASC_PROSTOR was changed
DREVESNICE	Integrated Administration and Control System (IACS)	Area of nurseries
DREV_SPR_M	Integrated Administration and Control System (IACS)	Month of change in the variable DREVESNICE
DREV_SPR_L	Integrated Administration and Control System (IACS)	Year of change in the variable DREVESNICE
DREV_RAZ_ID	Integrated Administration and Control System (IACS)	The code of the survey where the variable DREVESNICE was changed
ZACIMBE	Integrated Administration and Control System (IACS)	Area of aromatic and medicinal plants
ZAC_SPR_M	Integrated Administration and Control System (IACS)	Month of change in the variable ZACIMBE
ZAC_SPR_L	Integrated Administration and Control System (IACS)	Year of change in the variable ZACIMBE
ZAC_RAZ_ID	Integrated Administration and Control System (IACS)	The code of the survey where the variable ZACIMBE was changed
EKST_SAD	Register of Agricultural Holdings (RAH)	Area of extensive orchards
JAGODE	Register of Agricultural	Area of strawberries

	Holdings (RAH)	
JAGODE_SPR_M	Register of Agricultural Holdings (RAH)	Month of change in the variable JAGODE
JAGODE_SPR_L	Register of Agricultural Holdings (RAH)	Year of change in the variable JAGODE
JAGODE_RAZ_ID	Register of Agricultural Holdings (RAH)	The code of the survey where the variable JAGODE was changed
SEMENA_SADIKE	Integrated Administration and Control System (IACS)	Area of seed and seedlings
SEM_SAD_SPR_M	Integrated Administration and Control System (IACS)	Month of change in the variable SEMENA_SADIKE
SEM_SAD_SPR_L	Integrated Administration and Control System (IACS)	Year of change in the variable SEMENA_SADIKE
SEM_SAD_RAZ_ID	Integrated Administration and Control System (IACS)	The code of the survey where the variable SEMENA_SADIKE was changed
POLZI	Integrated Administration and Control System (IACS)	The amount of snails (in kg)
POLZI_SPR_M	Integrated Administration and Control System (IACS)	Month of change in the variable POLZI
POLZI_SPR_L	Integrated Administration and Control System (IACS)	Year of change in the variable POLZI
POLZI_RAZ_ID	Integrated Administration and Control System (IACS)	The code of the survey where the variable POLZI was changed
KZU_SPR_M	Integrated Administration and Control System (IACS)	Month of change in the variable KME_ZEM_UPO
KZU_SPR_L	Integrated Administration and Control System (IACS)	Year of change in the variable KME_ZEM_UPO
KZU_RAZ_ID	Integrated Administration and Control System (IACS)	The code of the survey where the variable KME_ZEM_UPO was changed

3.1.6 OBJECTIVE 6:

Review of agricultural holdings in the Statistical Farm Register:
- "cleaning" duplicate records,
- searching for links to administrative registers, where the "connection key" doesn't already exists,
- searching information of possible new agricultural holdings (horticultural producers).

Examined were agricultural holdings in the Statistical Farm Register (SFR) and duplicate records were discovered through different variables (personal ID, name, surname, phone number, address, farm address). 80 duplicates were discovered, and the status of agricultural holding has changed from eligible to non-eligible.

Based on the connection with the administrative Register of Agricultural Holdings (RAH), managed by the Ministry of Agriculture, Forestry and Food (MKGP), we were able to link our "not yet linked agricultural holdings" with their register. The direct linkage is important, since the administrative register has a legal basis behind it (Rules on the Register of Agricultural Holdings (OJ RS, No. 83/16, 23/17, 69/17, 72/18 in 35/19)) and it is updated by the farmers who want to apply for subsidies or want to raise animals, etc. The linkage was made based on different variables (personal ID, name, surname, phone number, address, farm address). In some 500 cases administrative ID (KMG_MID (SIF_MKG in SFR)) was assigned for the farms in the Statistical Farm Register. The current situation in the Statistical Farm Register is just 1,000 agricultural holdings with no connection to the administrative register (SIF_MKG is missing). Some of those agricultural holdings have a significant volume of production (for unknown reason they do not apply for subsidies), but most of them do not. Still all of them are considered as agricultural holdings above national threshold for agricultural production and are therefore treated as eligible farms.

Eligible agricultural holdings are those having the physical production above certain criteria:

at least one hectare of utilised agricultural area, or
less than 1 hectare of utilised agricultural area, but:

- at least 0.1 hectare of utilised agricultural area and 0.9 hectare of forest, or
- at least 0.3 hectares of vineyards and/or orchards, or
- two or more livestock units (LSU), or
- 0.15 to 0.3 hectare of vineyards/orchards and 1 or 2 LSU, or
- more than 50 beehives, or
- are market producers of vegetables, herbs, strawberries, mushrooms, flowers or ornamental plants.

We searched for new agricultural holdings on the internet. We assumed that some vegetable/flower producers might not be included in administrative farm register or Statistical Farm Register (SFR). The reason for that could be that the subsidies are payed per area and for vegetable/flower production you do not need a big area, but a production value is still big. Therefore we checked those, who have web pages built on the internet. We searched with key words, like: "horticultural producer, "vegetable producer", "selling vegetables, "selling flowers", etc. Those found with certain indication of production were manually linked with the Statistical Farm Register (name, surname address, etc.) and if not found, added into SFR. Based on information collected from the internet, we added some 10 producers to the Statistical Farm Register (SFR).

3.2 Converting the administrative farm ID (KMG-MID) to a statistical farm ID

3.2.1 OBJECTIVE 1

Preparation of the procedures for translating all administrative farm IDs into statistical farm IDs

For the purpose of "translation" of administrative farm IDs (KMG_MID) into statistical farm IDs (KME_ID), a new information source was established [SRKG_PREV@SQLETL1]. It contains a view table on the Statistical Farm Register (SFR) table "ADRESARJI", where only two variables are taken (first on is "KME_ID" – statistical identifier of agricultural holding and SIF_MKG (KMG_MID; administrative identifier of agricultural holding). The view table has a condition that takes only farms where variable "SIF_MKG" is not empty (we need only farms that are in the administrative register), and also removed are the duplicates in the Statistical Farm Register (SFR). Duplicate agricultural holdings are those that have in variable "ODG_ID"=5 – agricultural holding is a duplicate. That way we derived a "one to one" list of agricultural holdings. Access to this information source is restricted and every access is automatically recorded. When connecting to the information source "Statistical Farm Register – [SRKG_PREV@SQLETL1]", access is recorded in a special log file; MSSQL\Log.

Since new farms are created every day, the automatic update of the Statistical Farm Register on a daily basis is needed. For that reason every day at 17h the new farms in the Register of Agricultural Holdings (RAH) are automatically inserted into the Statistical Farm Register (SFR). All the new farms are for the time being treated as non-eligible, since the main purpose of daily update is to get a "fresh" list of newly added agricultural holdings for "translation" of administrative farm IDs (KMG_MID) into statistical farm IDs (KME_ID)). The newly added agricultural holdings get updated all the variables that are listed in "*Table 3.3. Variables updated with "automatic" updated in the Statistical Farm Register (SFR); table ADRESARJI".* With exception of variable ODG_ID, where all the agricultural holdings get the response ID as non-eligible. When the Statistical Farm Register (SFR) is regularly updated (twice per year) the variable "ODG_ID" – response ID is changed accordingly (see "objective 5" of "3.1 Modification of the Statistical Farm Register (SFR)").

New administrative sources can come to the Statistical Office on a daily basis. As mentioned in "objective 4" of "3.1 Modification of the Statistical Farm Register (SFR)", the Integrated Administration and Control System (IACS) is received by the Statistical Office as a pre-prepared dataset through the Secured File Transfer Protocol (SFTP) once a year and is used for updating the Statistical Farm Register (SFR) in September/October. When it reaches the single entry point (department involved with receiving the administrative data at SURS), the personal administrative ID (KMG_ID) needs to be translated.

For that purpose we created a "translator" – program that translates all the administrative KMG_MID (SIF_MKG in the Statistical Farm Register (SFR)) into KME_ID Unique statistical identifier of agricultural holding (See picture 3.2). The code for the "translator" was made in Visual Studio 2017, the tool SSIS (SQL Server Integration Services) with the help of computer language C#.

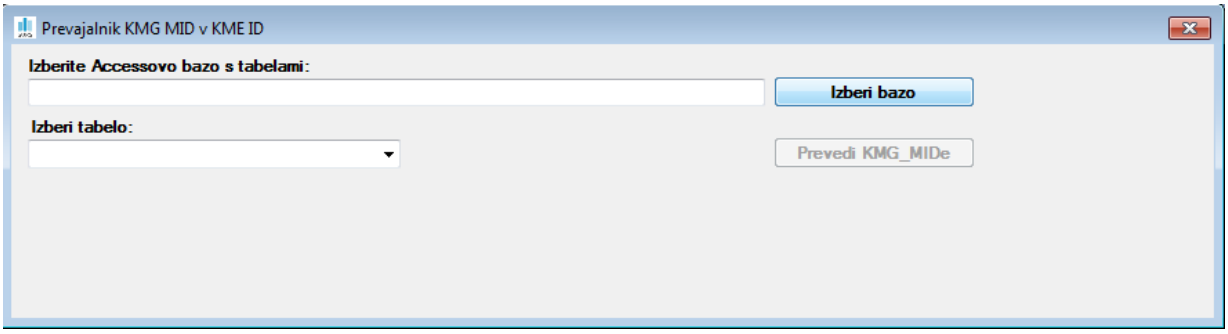
The program has two conditions to work:

1. It can read only files “accdb” or “mdb” (Access program). This was agreed with all administrative stakeholders that provide administrative data in pre-prepared datasets.
2. The administrative ID (KMG_MID) has to be in “text” format with the length 9 and the variable name has to begin with exact wording “KMG_MID”; example “KMG_MID” or “KMG_MIDxxxx”, where “xxxx” is optional text.

In the first combobox you chose the file you want to translate and in the second combobox you select the sheet you want to translate. When the sheet is selected, the program analyses the sheet and returns information if there are “KMG_MIDxxxx” variables in the sheet. If yes, then with the push of a button “Prevedi KMG_MIDe” the program executes the translation. It replaces all the administrative IDs (KMG_MID) with statistical IDs (KME_ID). Also the variable name is changed accordingly (from “KMG_MIDxxxx” into “KME_IDxxxx”). If there are some KMG_MID values that are not in the information source “Statistical Farm Register – [SRKG_PREV@SQLETL1]” (meaning that the KMG_MID values are not valid), then the first three numbers are replaced with a letter “A”. Example: “100123456” ==> “A123456”. That way those “agricultural holdings” can be easily found and examined.

If there are no “KMG_MIDxxxx” variables in the sheet, the program returns the text »In the table there are no “KMG_MIDxxxx” variables«: in this case the translation cannot be executed.

Picture 3.2. The “translator”; program that translates administrative ID (KMG_MID) into statistical ID (KME_ID)



3.2.2 OBJECTIVE 2

Determine the needs of translating historical data and preparing a timeline for all agricultural surveys containing the administrative farm ID

As stated in “objective 1” of “3.1 Modification of the Statistical Farm Register (SFR)”, the administrative farm ID (KMG_MID) is treated as a personal identifier and as such needs to be removed from all the working files and databases.

There were several meetings where all databases and files containing the administrative farm ID (KMG_MID) were located. It was decided that each employee in the agricultural department “cleans” its own files or writes instructions to the IT department on what and where the translation needed to be performed.

The deadline in performing all “translation” was 31 December 2018, so there was about two months’ time to perform all translation of the files to the statistical farm ID.

3.2.3 OBJECTIVE 3

Translation of the administrative farm ID into a statistical farm ID across all domains and in the Statistical Farm Register (SFR)

Since the Statistical Farm Register (SFR) is a confirmed secured personal database with all the names/surnames, personal IDs, addresses and administrative farm IDs, the translation of administrative ID (KMG_MID) was not needed in the Statistical Farm Register (SFR).

When examining the files, microdata usually contained the data on the administrative farm ID (KMG_MID). Since the microdata contained also the statistical farm ID (KME_ID), only the administrative farm ID (KMG_MID) was deleted and the file was clear from the perspective of the personal data protection. In other cases where microdata contained only the administrative farm ID (KMG_MID) the information source “Statistical Farm Register – [SRKG_PREV@SQLETL1]” was used to perform the translation into the statistical farm ID (KME_ID). At the end the administrative farm ID (KMG_MID) was deleted.

All “translations” of the administrative farm ID into a statistical farm ID was performed by 31 December 2018, and the administrative ID was deleted from all the databases and all files.

This task was also an opportunity to delete all the unnecessary working files which sometimes only provide confusion to workers.

3.3 Harmonization of input data

3.3.1 OBJECTIVE 1

Selection of variables important in various agricultural surveys (emphasis on the Farm Structure Survey)

Examined were three groups of variables:

1. Variables collected directly from agricultural holdings (questionnaire)
2. Pre-prepared variables
3. Derived variables

1. Variables collected directly from agricultural holdings (questionnaire)

Examined were the input variables in the agricultural census 2010, the Farm Structure Survey (FSS) 2013 and the Farm Structure Survey (FSS) 2016. The questionnaires are available on the internet (only in Slovenian):

Agricultural census 2010

Farm Structure Surveys (FSS) 2013

Farm Structure Surveys (FSS) 2016

The IDs of the variables are designed for each individual survey, so the comparison through years is not possible or at least very difficult.

2. Pre-prepared variables

The data for the Agricultural Census (AC) 2010 were collected as “computer assisted personal interview” (CAPI) and for the Farm Structure Survey (FSS) in 2013 and 2016 as “computer assisted telephone interview” (CATI). This means that some of the data for the agricultural holdings were not collected directly from farmers (not listed in the questionnaire), but were pre-prepared and put directly into the application for data collecting (also the database). Those variables are for example location IDs (village, municipality, region, etc.), statistical identifiers (person IDs, agricultural holding IDs, business IDs), code list for variable "NACIN_PRIDELAVE" – information on the type of farming (organic, conventional, integrated, farm in transition to organic farming, farm is partially organic farming, farm is partially in transition to organic farming), variables collected directly and only through administrative data (area of intensive orchards, hops, data on rural development, bovine animals (with exception of cows), etc.).

3. Derived variables

At the end usually some derived variables are prepared, which are in most cases used for easier dissemination purposes and are calculated based on some pre-prepared definitions (standard output of agricultural holding, typology of agricultural holding, livestock units, annual working units for whole agricultural holding, etc).

3.3.2 OBJECTIVE 2

Determination of the encryption method (standard for all agricultural surveys)

When examining the codes and making comparisons between surveys, we concluded that the main purpose of the cross-code list should be to provide a comparison through years and a comparison with other surveys. We analysed how to create codes for the purpose of tabulation, but since the data are so complex and since there are many variables that are a subcategory of another one, this idea was put aside. We concluded that a part of the code (first few letters) need to be established so that some main tabulation could be performed (cattle - total, pigs – total, arable land – total, vineyard – total, etc.).

The totals in the questionnaire did not get any codes, since that way we would get duplicates when making a tabulation. For example pigs – total does not get any code, but pigs total can easily be summed, when taking only a part of the code (see the example below).

Based on experiences of other countries, we decided to make a code system that can be used now and also in the future when new variables would be added into the survey. For that purpose we constructed a multi-part code with the length of 10 places.

Example; meaning of the CODE A01_03_007 - Breeding boars:

A – Domain (Agricultural domain; we leave room for other statistical domains to approach to our encryption method if suited):

_ – A mark that divides the multi-part code for easier examination and division of the code.

01 – Main theme (Livestock; we added two places to divide the agricultural domain into different aspects of data collecting (others are land use, irrigation, etc.).

03 – Subtheme (Pigs: we added two places for division of the livestock on species. To this place (first six places) the codes can be tabulated and the data are comparable through years).

007 – Variables asked to agricultural holdings or variables in the database (Breeding boars)

Keeping a lot of “empty“ places in each part of the code allows us to ask new variables in the future and still use the same encryption method.

For some other variables we decided not to use a multi-part code since the variables are more stable and will not change through years:

For example;

Statistical business ID ➔ “SIR”

Standard output ➔ “SO_KMET_EUR”

Annual working units ➔ “PDM_KM”

Less favourable area ➔ “OMD_KMG”

Year of birth of the manager ➔ “U_ROJEN”

3.3.3 OBJECTIVE 3

Implementation in the Farm Structure Surveys (2010, 2013 and 2016)

A cross-code list was established (in Excel program) by using the name of the variables that were used in the Farm Structure Surveys (2010, 2013 and 2016). The codes there were used for better guidance through the questionnaire (B1, B2, B3, etc.; codes also used in the database). Added were descriptions and the location where an individual variable is located in the database. A proper multi-part code was assigned to each variable.

Where the code will also be used in the agricultural census 2020, the code is written in the table (see Annex 2).

3.3.4 OBJECTIVE 4

Technical harmonization of databases - implementation for the agricultural census 2020; implementation in other surveys after 2020

When the cross-code list was established, the SAS program was used to collect the information from the code list and all the tables. Established was a harmonized table with all the data from the previous Farm Structure Surveys (2010, 2013 and 2016).

Established was a new information source [AC2020@SURSOLAP] (See objective 6 of "3.4 Preparation for the agricultural census 2020"), where a link table was placed in order to have the data available for controls or to use them for imputation purposes.

Technical implementation of the codes was made for the CORE agricultural census questionnaire, which is scheduled to be performed in June 2020 (will be published on SURS's website just before the data collection – June 2020). Modules for the Integrated Farm Survey are scheduled to be performed in December 2020 and the questionnaire and codification are in development. Since the rules for the codification are set, the implementation for the new variables can be made rapidly.

3.3.5 OBJECTIVE 5

Study visit

Visited was the Estonian statistical office (21–22 February 2019), where we gained knowledge on their method for encryption of the input variables and the whole process of conducting the agricultural census, for which we are very grateful. The whole study visit was set up in the way that representatives of sectors and departments presented a comprehensive overview of their field of work. A fruitful debate on specific items of each topic took place during individual presentations. Details of the whole study visit are in Annex 1.

3.4 Preparation for the agricultural census 2020

3.4.1 OBJECTIVE 1

Reviewing the new Integrated Farm Statistics (IFS) Regulation requirements and defining the data source for each variable in the new IFS Regulation

On 18 July the new regulation on Integrated Farm Statistics (IFS) was published. "*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" ("IFS regulation").* It is a legal basis for conducting the Agricultural Census (AC) 2020.

On 29 November also the implementing regulation was published, where all the variables and definitions were described: "*COMMISSION IMPLEMENTING REGULATION (EU) 2018/1874 of 29 November 2018 on the data to be provided for 2020 under Regulation (EU) 2018/1091 of the European Parliament and of the Council on integrated farm statistics and repealing Regulations (EC) No 1166/2008 and (EU) No 1337/2011, as regards the list of variables and their description".*

The new regulation and the implementing regulation were analysed and the data sources for the variables were defined. The agricultural census in Slovenia will be executed with all three main sources of data, which are specified in Article 4 of the IFS regulation:

(a) statistical surveys,

(b) administrative data sources,

(c) other sources, methods or innovative approaches.

The statistical survey will be conducted on the farms that are not in administrative sources (for CORE variables) and also for the farms that will be included in the so-called Module surveys (See Article 7 of the "IFS regulation").

The administrative sources will be used in most of the CORE variables and also some other variables as auxiliary data source.

The innovative methods will be used in some part of the variables (See objective 2 of "3.4 Preparation for the agricultural census 2020").

The whole list of data sources for agricultural census variables was constructed and sent to Eurostat as stated in article 4, paragraph 3 of the "IFS regulation" by the end of 2019.

3.4.2 OBJECTIVE 2

Defining variables for which we do not have administrative data and considering the possibility of model estimation

Variables for which we do not have a direct linkage to administrative sources and which have to be collected on the whole population of agricultural holdings (so-called CORE variables) were identified and model solutions are in development/developed. Besides 'traditional' statistical modelling methods, also geospatial modelling was utilized which enabled a more thorough exploitation of data sources (using spatial component of the data sources).

Already in the past, statistics relied heavily on administrative (IACS data). Despite its reliability and adequate coverage, not all information is explicitly captured in a single data source. For those cases, survey questions had to be utilized in the past.

In the scope of this project, modelling processes were developed in order to extract desired information by combining multiple data sources. Geospatial modelling was utilized which enabled a more thorough exploitation data sources (using spatial component of the data sources).

Developed models enable us to make a full capture of a state for a reference date without creating any respondent burden. Processes are also reusable for other areas of statistics. Reuse in other countries depends on availability of comparable data sources. Reuse would probably require processes adaptation at best and at worst only processing logic would be applicable.

Since the geospatial modelling runs on basis of agricultural plots instead of agricultural holdings, the geolocation and the granularity of the obtained data is much finer. This opens possibilities for various data aggregations and novel indicators calculations. Results can be provided timely, given the timeliness of used data sources. In most cases the processing time is shorter compared to traditional data collection.

Main limitation of presented models is the quality of the administrative data utilized. Currently all sources are updated and of good quality, however, in case of a disturbed maintenance the quality could drop.

Land Parcel Identification System (LPIS)

The purpose of Land Parcel Identification System (LPIS) in Slovenia is to implement common agricultural policy of the European Union measures. LPIS is a part of the administrative farm register and works as a spatial representation of areas utilized by agricultural holdings. The reference parcel of LPIS is farmer's block (blok). A blok is a compact area of agricultural land in agricultural use by one agricultural holding (exceptions are defined in the legislation). The agricultural holding land use unit (agricultural plot) is a compact area of agricultural land with the same type of land use within each blok. Farmers declare their land in the form of agricultural plot. Published LPIS data is available at the agricultural plot level with block id as an attribute.

Land Cadastre

Land Cadastre is an official register of land parcels. Land Cadastre includes up-to-date information on parcels and includes the collection of documentation and information necessary for historical track changes of parcels. The collection of documentation (in analogue and digital form) includes files and other documents that were originally used for the official registration of parcels. The collection of documentation is stored permanently.

A basic Land Cadastre unit is the land parcel (hereinafter: parcel). It represents aggregate land that lies within one cadastral municipality and is recorded in the Land Cadastre with the border and marked with an identification mark.

The following information are kept in the Land Cadastre: the parcel identification code, border, surface, owner, manager, actual land use, land under the building, land benefit.

The relation to the Register of Spatial Units, Building Cadastre and Land Registry is also provided.

Private surveying companies are using this data in geodetic procedures. Individual municipal departments and state administration offices are using Land Cadastre data as a basis for the spatial planning, environmental protection, valuation and real estate taxation, management of transport networks and public infrastructure facilities, etc.

The coordinate system D96/TM is used.

The Land Use Database

Agricultural and Forestry Land Use Database is an official national database of agricultural and forestry land use in Slovenia. Basic element in the database is a land use polygon which represent the unique part of land with the same land use. Its primary purpose is to determine the land cover / land use in high scale to be applied as a control layer for implementation of measures of the common agricultural policy of the European Union.

Irrigable land register (Katmesina)

It is a national register of melioration systems (drainage and irrigation). Besides the geolocations and documentation on individual existing systems also data on planed systems is collected. The register is currently under revision and is being updated to the current state of systems. Completion is planned for summer 2020.

The land-rating value

The land-rating value of agricultural land is one of the most important records/inventories of agricultural land quality, on which spatial planning and agricultural land protection decisions are based, as well as land assignment in land consolidation processes. Land valuation was established state-wide in Slovenia in 2008 by a conversion of cadastral classification data into land-rating values.

The determination of values is described in Rules on determining and administering land rating (Official Gazette RS, št. 47/08).

It is calculated using the formula below:

$$B = \sqrt{T_x \times K_x \times R_x} \times \left(1 - \frac{\sum \%\ special\ effects_x}{100}\right)$$

Parameters:

B	land rating value
T	points for soil properties
K	points for climate properties
R	points for terrain properties
Σ %*special effects*	sum of shares of special effects

Points are calculated as follows:

- Points for soil properties are a numerical value assigned based on field examination and pedological properties for each land parcel. Pedological properties are based on bedrock, soil type, degree of development and texture. Points range is between 1 and 100.
- Points for climate properties are obtained based on data of the nearest meteorological station. Temperature, precipitation and microclimatic conditions are considered. Points range is between 1 and 10 (1 being least and 10 most appropriate for agriculture).
- Points for terrain are estimated based field examination (based on inclination, accessibility, gradient, energy…). Points range is between 1 and 10 (1 being least and 10 most appropriate for agriculture).

Processes

Main reference for agricultural census in Slovenia is LPIS (IACS) data frame. Its spatial reference is explicit but changing on yearly basis with its spatial units (agricultural plots) rarely matching with spatial units of other geospatial data sources. Thus, spatial intersection between the sources is necessary to obtain properties from additional data sources for the spatial units of interest (agricultural plot). This is done in ESRI ArcGIS Pro software environment. Additional processing depends on each individual procedure.

Rented UAA

An example of such workflow is described in the article “The Integration of Administrative Data for the Identification of the Ownership of Agricultural Land” (See Annex 3), which has been presented on ICAS 8 conference in New Delhi).

Market garden/Open field

Fresh vegetables (including melons) and strawberries:

- Fresh vegetables (including melons) and strawberries grown in rotation with horticultural crops (**market gardening**)
- Fresh vegetables (including melons) and strawberries grown in rotation with non-horticultural crops (**open field**)

LPIS of current and previous year overlay was created. Intersection areas were classified as shown in the table below. Main idea of the process is that areas, where fresh vegetables are grown in two consecutive years is ruled as market garden, whereas areas in rotation with non-horticulture crops is ruled as open field.

Fresh vegetables (including melons) and strawberries - *in **previous** year*	Fresh vegetables (including melons) and strawberries - *in **current** year*	Open field *or* Market gardening
Yes	Yes	market gardening

Yes	No	/
No	Yes	open field
No	No	/
/	Yes	open field
/	No	/

Rough grazing

The process is still under development. Thus far, potential data sources have been identified and tested (compared against historical data). As main input, the land-rating value has been tested, as it indicates soil, terrain and climate properties of each agricultural plot. This accounts for the natural potential for use. It was recognized, however, that also anthropogenic limitations must be considered. Therefore additional data sources have been looked for, namely Natura 2000 (Metadata) and List of less favoured areas (Official Gazette RS, št. 27/03 in 116/04), etc. Testing proved that finer granularity will be required, since both forces have a large spatial coverage thus overestimating rough grazing. Subdividing of existing data sources and identification of additional ones is currently under process.

Irrigable land

The geospatial source of Irrigable land register has been identified as suitable for use for agricultural census. The procedures for the processing are prepared, however testing was not possible with the data source under revision. As soon as the revision is complete the data will be adopted and processed.

Farm land structure

To identify land characteristics, 3 data sources will be combined as shown in the table below. It provides a full coverage with explicit spatial dimension.

Land characteristics		**Source**
Utilised agricultural area (UAA)		LPIS
Other farmland		
-	Unutilised agricultural land	The Land Use Database
-	Wooded area	The Land Use Database
-	Other land (land occupied by buildings, farmyards, tracks, ponds and other non-productive areas)	Land Cadaster

3.4.3 OBJECTIVE 3

An overview of administrative data sources and the identification of possible new ones

An overview of all administrative data already gathered by the statistical office was examined and the possibility of new data sources was analysed. For that purpose a lot of phone calls and meetings with data providers were made and the results were encouraging.

The data providers show a lot of interest, so that the data already collected for some administrative purpose are also used for statistical needs and not to bother the reporting units twice. This was also an opportunity to change the data transmission method from a CD to more secure and fast way. Now administrative data come to the statistical office only in two ways:

- Through a Secured File Transfer Protocol (SFTP) server
- Direct access to the administrative database where tables that are needed for certain use can be downloaded

Administrative data that were gathered and general agreements and technical protocols that were signed:

- Register of Agricultural Holdings (RAH), managed by the Ministry of Agriculture, Forestry and Food (MKGP) (integrated are data on holder/manager, orchards, strawberries, vineyard, olive trees, area under greenhouse)
- Register of Holders of Farmed Animals, managed by the Administration for Food Safety, Veterinary Sector and Plant Protection (data on beehives, cattle, pigs, sheep and goats, equidae)
- Integrated Administration and Control System (IACS) data, managed by the Agency for Agricultural Markets and Rural Development (data on land parcel area by different crop and number of livestock on agricultural holding, data on manure treatment, data on animal housing, etc.)
- Rural development data, managed by the Agency for Agricultural Markets and Rural Development
- Data from “Rules on records for the milk sector and on the market information system for the milk and milk products market”, managed by the Agency for Agricultural Markets and Rural Development (data on milk production at the farm level)
- Central Cattle Database, managed by the Agricultural Institute of Slovenia (data on cattle – milking cows)
- Organic Farming Register, managed by the Ministry of Agriculture, Forestry and Food (MKGP) (data on organic land area and the number of animals)
- Register of Irrigation Systems, managed by the Ministry of Agriculture, Forestry and Food (MKGP)
- Register of Education and Training for the Needs of Agriculture and Rural Development, managed by the Ministry of Agriculture, Forestry and Food (MKGP) (data on education and training of the holder/manager)
- Data on hops production, managed by the Ministry of Agriculture, Forestry and Food (MKGP)
- Data on cannabis and garden poppy production, managed by the Ministry of Agriculture, Forestry and Food (MKGP)
- Data on rented state land, held by the Farmland and Forest Fund of the Republic of Slovenia (data on the land that is rented out to agricultural holdings)
- Data on production of agricultural plant seed, managed by the Administration for Food Safety, Veterinary Sector and Plant Protection (certified seed production area)

3.4.4 OBJECTIVE 4

Updating and preparing new agreements (technical protocols) on the transmission of administrative data

In general agreements the following themes were agreed upon:

- Who will use the data (Statistical Office of the Republic of Slovenia (SURS) and holder of the administrative data)
- What data are the subject of the agreement
- Legal basis on which the data were gathered and the basis for the statistical office to use them
- For details on concrete data and data transmission the technical protocols will be established
- The methodological explanations will be made available if needed
- If the data are sent to the European Commission, the statistical office will be notified
- No financial obligations for both parties

In technical protocols the following themes were agreed upon:

- Who will use the data (Statistical Office of the Republic of Slovenia (SURS) and holder of the administrative data)
- General description of the data
- What is prepared and when
- Name of the tables
- Technical administrator and subject administrator for both parties (name/surname and email of persons)
- List of variables and their description
- Transfer of the data and data protection details
- Deadlines for sending the data

Signed were all general agreements and technical protocols for all administrative data, listed in objective 3 of the “3.4 Preparation for the agricultural census 2020”.

At the Statistical Office of the Republic of Slovenia (SURS) it was agreed that in the future themes that are in the general agreement will be moved to technical protocols so that only one document will be prepared.

3.4.5 OBJECTIVE 5

Preparing documentation for each individual administrative source

Documentation for each individual administrative source was included in general agreements and technical protocols (legal basis, list of variables and their description, etc.) (See objective 3 of the “3.4 Preparation for the agricultural census 2020”).

The only exception with exact documentation is the Register of Agricultural Holdings (RAH), to which the statistical office has direct access. It contains many tables that need to be linked (combined) to get the necessary data. The documentation was organized at the statistical

office, and it was agreed with the Ministry (added in the technical report) that if the statistical office needs an additional explanation of the variables, the support would be needed.

3.4.6 OBJECTIVE 6

Preparing a database with all administrative sources and other auxiliary statistical sources in order to perform the analysis of the quality of the obtained administrative data sources (comparing administrative data with the existing-historical statistical surveys)

A test SQL information source [AC2020@SURSOLAP] was established and main administrative sources were entered into the database (listed in objective 3 of "3.4. Preparation for the agricultural census 2020"). The administrative data were examined and test tabulation was made. Since also in previous surveys administrative data were used, the tabulation and the comparison with previous data is encouraging.

The quality of all administrative data was assessed when running the analysis and making technical protocols.

3.4.7 OBJECTIVE 7

Conducting study visits in countries with experiences in the field of administrative agricultural census

We participated in a workshop that was organized by the statistical office of Poland. It was organized in the scope of the grant for modernization of agricultural statistics. The main topic was the agricultural census 2020 and the ongoing preparation activities connected to it. It was organized in presentation format held by country representatives. There was open room for discussion during the presentations and for networking during the brakes (See Annex 4).

We also gained knowledge on the whole process of conducting the agricultural census from colleagues from the Estonian statistical office (See Annex 1).

List of abbreviations

Acronym	Description
UAA	Utilized agricultural area
SFR	Statistical Farm Register
RAH	Register of Agricultural Holdings
MKGP	Ministry of Agriculture, Forestry and Food
SURS	Statistical Office of the Republic of Slovenia
KMG_MID	Unique administrative identifier of agricultural holding used by the MKGP
KME_ID	Unique statistical identifier of agricultural holding used by SURS
FADN	Farm Accountancy Data Network
IACS	Integrated Administration and Control System
SFTP	Secured File Transfer Protocol
CAPI	Computer Assisted Personal Interview
CATI	Computer Assisted Telephone Interview
LPIS	Land Parcel Identification System
IFS regulation	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".

Annex 1. Study visit Estonia

The study visit regarding Eurostat grants for 2017; Objective: 08.4 – Provide quality agriculture, fisheries and forestry statistics; Module (DTM): 08.4.11 Agricultural statistics; Title of action: Modernization of agricultural statistics.

Dates: 21–22 February 2019

Location: Statistics Estonia,Tallinn

Participants from Slovenia: Statistical Office of the Republic of Slovenia – Agriculture, Forestry, Fishing and Hunting Statistics Section

1. Summary

Statistics Estonia hosted us on a two-day visit, for which we are very grateful, as their system is very advanced and we learned and exchanged a lot of information.

The main topics of the study visit were gathering the experience on the method for encryption of the input variables and discussion on the agricultural census 2020. The whole visit was set up so that representatives of sectors and departments presented a comprehensive overview of the fields of work and the organization of their statistical office. An individual debate on specific interests of the Statistical Office of the Republic of Slovenia (SURS) and a comparison of the systems of both organizations took place during individual presentations.

The information obtained during the study visit will facilitate work on the project of modernization of agricultural statistics in Slovenia. The items presented will improve the results of the project, and consequently the agricultural census itself.

2. Content

- The new Integrated Farm Statistics (IFS regulation) will cause quite a lot of coordination in the whole statistical office and in consequence we will have to become more flexible when gathering new information.
- In Estonia the response through computer assisted web interview (CAWI) is somewhere from 20% to 50%, which is quite high. The CAWI for the Farm Structure Survey in Slovenia was never developed. This is something that could be built in the future to reduce the response burden to some farmers ➔ the plan for the Slovenian Agricultural Census 2020 is administrative CORE census, so the reduce the response burden would be anyhow significant.
- The higher threshold given in the IFS regulation will cause that a lot of agricultural holdings will be out of scope, so they won't be covered. Based on national needs, each country will decide what they will do (either to make frame extension or not).
- Since the MODULES for the census can be made on a sample, the dissemination on the variables in CORE and MODULES will cause some discrepancies. The solution to that issue has not yet been found.
- Data on own and rented land in Estonia are obtained directly from the administrative register, since agricultural holdings must provide information on rented land when they apply for subsidies. In Slovenia such information does not exist. In our case we plan to get the area of rented land from the Land Parcel Identification System (LPIS) in combination with the Land Register (information on ownership of individual parcel).
- A discussion on grant application was made and we agreed that a lot of costs would be fixed costs due to new regulation's requirements.

- Administrative data in Estonia are very well organized and all the agreements with reporting institutions are maintained regularly. There is a special department dealing only with administrative sources. In Slovenia administrative data are dealt with in the department that uses a specific administrative source.
- Imputations in Estonia are made in a special department, and the transfer from the Agricultural Department to the "imputation" department is made by the third person, so communication between all participants is quite a challenge.
- In Estonia the system of metadata (iMeta) is well organized and harmonized.
    - It is a single system for the whole statistical office.
    - The encryption method of the input variables is not harmonized with Eurofarm variables.
    - When a new requirement for a variable arises, this variable is placed into iMeta and from there it is used for the statistical survey.
    - From iMeta annual program for statistical surveys is drawn.
    - There is a direct connection to micro data (location to the server where the data are located).
    - Described are the types of variables: categorical (link to code list), numerical and textual.
    - The names of the variables remain the same over the years (for easier work with time series).
    - The names of the variables are built based on the description of the variable (for example "Bovine – total" would be named BOV_TOT). They do not have systematically encrypted names.
    - iMeta has input, output and also derived variables. Using this principle there is no mix-up with dissemination. Also if the name of the input and output variable is the same, they are inserted twice in the iMeta.
- The purpose and maintenance of the statistical farm register was discussed and experiences were exchanged.

3. Agenda

Thursday, 21.02. 2019		
Time (08.30-16.30)	Activity	Short description
08.30 – 9.30	*Welcome; Introduction of the participants*	
9.30 – 12.00	*Overview of the activities in IFS 2020*	Presentation of the national Farm Structure Survey (FSS) in Estonia 2016 and discussion on the coming agricultural census in 2020.
12.00 – 15.00	*Use of administrative sources*	A discussion on the use of administrative data in the FSS (IFS) and possible issues when combining different ones into one FSS database. Discussion on agreements and technical protocols with administrative data providers.
15.00 – 16.30	*Discussion on the new IFS legislation (2018/109)*[1]	Discussion on the differences between the old and new legislation and the effect that it will have on the AC 2020 data (possible issues with coverage?, reference date?, etc.)

Friday, 22.02. 2017		
Time (08.30-16.30)	Activity	Short description
08.30 – 12.00	*Harmonization of input data*	Presentation on the work done in harmonizing the input and output variables in the metadata system. Explanation on how the metadata system works, how the code is built and for which other purposes it is used (dissemination?, administration?, etc.)
12.00 – 13.30	Imputations and model estimations	Discussion on the activities done on imputations and model estimation.
13.30 – 15.00	Statistical Farm Register	Discussion on the purpose and maintenance of the statistical farm register (what is the use, how it is built, updating, variables, etc.)
15.00 – 16.30	*General discussion*	

Annex 2. Code list for the Agricultural Census 2010, the Farm Structure Survey 2013, 2016 and the Agricultural Census 2020

CODE	FORMAT CODE	VARIABLE 2010	VARIABLE 2013	VARIABLE 2016	VARIABLE 2020
KME_ID	7 CHAR	KME_ID	KME_ID	KME_ID	KME_ID
ZAP_ST	6 NUM	ZAP_ST	ZAP_ST	ZAP_ST	
MAT_ST	10 CHAR	MAT_ST	MAT_ST	MAT_ST	MAT_ST
NAC_PRIDEL	1 CHAR	NAC_PRIDEL	NAC_PRIDEL	NAC_PRIDEL	
EPK_1997	1 CHAR	EPK_1997	EPK_1997	EPK_1997	
EPK_EU_IFS	1 CHAR				EPK_EU_IFS
EPK_SI_2020	1 CHAR				EPK_SI_2020
SO_KMET_EUR	14.4 NUM	SO_KMET_EUR	SO_KMET_EUR	SO_KMET_EUR	
PDM_KG	11.5 NUM	PDM_KG	PDM_KG	PDM_KG	
TIP_KMETIJE_2010	3 CHAR	TIP_KMETIJE_2010	TIP_KMETIJE_2010	TIP_KMETIJE_2010	
RAZISKAVA	25 CHAR	RAZISKAVA	RAZISKAVA	RAZISKAVA	
MESEC_RAZ	2 CHAR	MESEC_RAZ	MESEC_RAZ	MESEC_RAZ	
LETO_RAZ	4 CHAR	LETO_RAZ	LETO_RAZ	LETO_RAZ	
GVZ_GOVEDO	10.4 NUM	GVZ_GOVEDO	GVZ_GOVEDO	GVZ_GOVEDO	GVZ_GOVEDO
GVZ_PRASICI	10.4 NUM	GVZ_PRASICI	GVZ_PRASICI	GVZ_PRASICI	GVZ_PRASICI
GVZ_DROBNICA	10.4 NUM	GVZ_DROBNICA	GVZ_DROBNICA	GVZ_DROBNICA	GVZ_DROBNICA
GVZ_PERUTNINA	10.4 NUM	GVZ_PERUTNINA	GVZ_PERUTNINA	GVZ_PERUTNINA	GVZ_PERUTNINA
GVZ_KUNCI	10.4 NUM	GVZ_KUNCI	GVZ_KUNCI	GVZ_KUNCI	GVZ_KUNCI
GVZ_KONJI	10.4 NUM	GVZ_KONJI	GVZ_KONJI	GVZ_KONJI	GVZ_KONJI
GVZ_DRUGA	10.4 NUM	GVZ_DRUGA	GVZ_DRUGA	GVZ_DRUGA	GVZ_DRUGA
GVZ_SKUPAJ	10.4 NUM	GVZ_SKUPAJ	GVZ_SKUPAJ	GVZ_SKUPAJ	GVZ_SKUPAJ
MID_HIS_ST_KMG	8 NUM	MID_HIS_ST_KMG	MID_HIS_ST_KMG	MID_HIS_ST_KMG	MID_HIS_ST_KMG
NA_ID_KMG	3 CHAR				NA_ID_KMG
NA_MID_KMG	8 CHAR				NA_MID_KMG
NASELJE_KMG	40 CHAR				NASELJE_KMG
OB_ID_KMG	3 CHAR				OB_ID_KMG
OB_MID_KMG	8 CHAR				OB_MID_KMG
OBCINA_KMG	40 CHAR				OBCINA_KMG
KO_MID_KMG	8 NUM	KO_MID_KMG	KO_MID_KMG	KO_MID_KMG	
KO_ID_KMG	4 NUM	KO_ID_KMG	KO_ID_KMG	KO_ID_KMG	
KO_IME_KMG	40 CHAR	KO_IME_KMG	KO_IME_KMG	KO_IME_KMG	
CENILNI_OKOLIS_KMG	5 CHAR	CENILNI_OKOLIS_KMG	CENILNI_OKOLIS_KMG	CENILNI_OKOLIS_KMG	
UE_ID_KMG	5 NUM	UE_ID_KMG	UE_ID_KMG	UE_ID_KMG	
UE_IME_KMG	40 CHAR	UE_IME_KMG	UE_IME_KMG	UE_IME_KMG	
PROSTORSKI_OK_MID_KMG	8 NUM	PROSTORSKI_OK_MID_KMG	PROSTORSKI_OK_MID_KMG	PROSTORSKI_OK_MID_KMG	
GEO_SIRINA_KMG	6.4 NUM	GEO_SIRINA_KMG	GEO_SIRINA_KMG	GEO_SIRINA_KMG	
GEO_DOLZINA_KMG	6.4 NUM	GEO_DOLZINA_KMG	GEO_DOLZINA_KMG	GEO_DOLZINA_KMG	
OMD_KMG	1 CHAR	OMD_KMG	OMD_KMG	OMD_KMG	
PREJEMNIK_IAKS	1 CHAR				PREJEMNIK_IAKS
MLADI_KMET	1 CHAR				MLADI_KMET
EFA_KMG	9 NUM			EFA_KMG	
U_ROJEM	4 NUM				U_ROJEN
U_SPOL	1 CHAR				U_SPOL
U_DELO	3 CHAR				U_DELO
U_LETO_POSTAL_U	4 NUM				U_LETO_POSTAL_U
U_IZOBRAZBA	1 CHAR				U_IZOBRAZBA
U_USPOSABLJANJE	1 CHAR				U_USPOSABLJANJE
PRP_71010	1 CHAR	PRP_71010	PRP_71010		
PRP_71020	1 CHAR	PRP_71020	PRP_71020		
PRP_71030	1 CHAR	PRP_71030	PRP_71030		
PRP_71040	1 CHAR	PRP_71040	PRP_71040		
PRP_71050	1 CHAR	PRP_71050	PRP_71050		
PRP_71060	1 CHAR	PRP_71060	PRP_71060		
PRP_71070	1 CHAR	PRP_71070	PRP_71070		
PRP_71080	1 CHAR	PRP_71080	PRP_71080		

PRP_71081	1 CHAR	PRP_71081	PRP_71081		
PRP_71090	1 CHAR	PRP_71090	PRP_71090		
PRP_71100	1 CHAR	PRP_71100	PRP_71100		
PRP_71110	1 CHAR	PRP_71110	PRP_71110		
PRP_01	1 CHAR			PRP_01	
PRP_02	1 CHAR			PRP_02	
PRP_03	1 CHAR			PRP_03	
PRP_04	1 CHAR			PRP_04	
PRP_05	1 CHAR			PRP_05	
PRP_06	1 CHAR			PRP_06	
PRP_07	1 CHAR			PRP_07	
PRP_08	1 CHAR			PRP_08	
PRP_09	1 CHAR			PRP_09	
PRP_10	1 CHAR			PRP_10	
PRP_11	1 CHAR			PRP_11	
PRP_12	1 CHAR			PRP_12	
PRP_13	1 CHAR			PRP_13	
PRP_14	1 CHAR			PRP_14	
PRP_15	1 CHAR			PRP_15	
PRP_16	1 CHAR			PRP_16	
PRP_17	1 CHAR			PRP_17	
SVETOVALNI_OK_KMG	8 NUM		SVETOVALNI_OK_KMG	SVETOVALNI_OK_KMG	
KMETIJSKI_ZAVOD_ID_KMG	8 NUM		KMETIJSKI_ZAVOD_ID_KMG	KMETIJSKI_ZAVOD_ID_KMG	
KMETIJSKI_ZAVOD_IME_KMG	40 CHAR		KMETIJSKI_ZAVOD_IME_KMG	KMETIJSKI_ZAVOD_IME_KMG	
STRATUM	1 NUM		STRATUM	STRATUM	
UTEZ	17.15 NUM		UTEZ	UTEZ	
NUTS3_ID_KMG	3 NUM	NUTS3_2013_ID_KMG	NUTS3_2013_ID_KMG	REGIJA_ID_KMG	NUTS3_ID_KMG
NUTS3_IME_KMG	40 CHAR	NUTS3_2013_IME_KMG	NUTS3_2013_IME_KMG	REGIJA_IME_KMG	NUTS3_IME_KMG
NUTS2_ID_KMG	1 NUM	NUTS2_2013_ID_KMG	NUTS2_2013_ID_KMG	NUTS2_ID_KMG	NUTS2_ID_KMG
NUTS2_IME_KMG	40 CHAR	NUTS2_2013_IME_KMG	NUTS2_2013_IME_KMG	NUTS2_IME_KMG	NUTS2_IME_KMG
NOVI_SID_EMSO	9 NUM	NOVI_SID_EMSO			NOVI_SID_EMSO
A01_01_001	8 NUM	B56	B50	B50	A01_01_001
A01_01_002	8 NUM	B57	B51	B51	
A01_01_003	8 NUM	B58	B52	B52	
A01_01_004	8 NUM	B59	B53	B53	
A01_01_005	8 NUM	B60	B54	B54	A01_01_005
A01_01_006	8 NUM	B61	B55	B55	
A01_01_007	8 NUM				A01_01_007
A01_01_008	8 NUM				A01_01_008
A01_02_001	8 NUM	GOV_BIKI_DO_1	GOV_BIKI_DO_1	GOV_BIKI_DO_1	A01_02_001
A01_02_002	8 NUM	GOV_TEL_DO_1	GOV_TEL_DO_1	GOV_TEL_DO_1	A01_02_002
A01_02_003	8 NUM	GOV_BIKI_1_2	GOV_BIKI_1_2	GOV_BIKI_1_2	A01_02_003
A01_02_004	8 NUM	GOV_TEL_1_2	GOV_TEL_1_2	GOV_TEL_1_2	A01_02_004
A01_02_005	8 NUM	GOV_BIKI_NAD_2	GOV_BIKI_NAD_2	GOV_BIKI_NAD_2	A01_02_005
A01_02_006	8 NUM	GOV_TEL_NAD_2	GOV_TEL_NAD_2	GOV_TEL_NAD_2	A01_02_006
A01_02_007	8 NUM	GOV_KRAVE_DOJILJE	GOV_KRAVE_DOJILJE	GOV_KRAVE_DOJILJE	A01_02_007
A01_02_008	8 NUM	GOV_KRAVE_MOLZNIC	GOV_KRAVE_MOLZNIC	GOV_KRAVE_MOLZNIC	A01_02_008
A01_03_001	8 NUM	B6	B6	B6	
A01_03_002	8 NUM	B7	B7	B7	
A01_03_003	8 NUM	B8	B8	B8	
A01_03_004	8 NUM	B9	B9	B9	
A01_03_005	8 NUM	B10	B10	B10	
A01_03_006	8 NUM	B11	B11	B11	
A01_03_007	8 NUM	B12	B12	B12	A01_03_007
A01_03_008	8 NUM	B13	B13	B13	
A01_03_009	8 NUM	B14	B14	B14	
A01_03_010	8 NUM	B15	B15	B15	
A01_03_011	8 NUM	B16	B16	B16	
A01_03_012	8 NUM				A01_03_012
A01_03_013	8 NUM				A01_03_013
A01_03_014	8 NUM				A01_03_014
A01_03_016	8 NUM				A01_03_016

A01_04_001	8 NUM	B38	B38	B38	A01_04_001
A01_04_002	8 NUM	B39			
A01_04_003	8 NUM	B40			
A01_04_004	8 NUM	B41			
A01_04_005	8 NUM	B42			
A01_04_006	8 NUM		B39	B39	A01_04_006
A01_04_007	8 NUM	B43	B40	B40	A01_04_007
A01_04_008	8 NUM	B44	B41	B41	
A01_05_001	8 NUM	B47	B44	B44	A01_05_001
A01_05_002	8 NUM	B48			
A01_05_003	8 NUM	B49			
A01_05_004	8 NUM	B50			A01_05_004
A01_05_005	8 NUM	B51			A01_05_005
A01_05_006	8 NUM		B45	B45	A01_05_006
A01_05_007	8 NUM	B52	B46	B46	A01_05_007
A01_05_008	8 NUM	B53	B47	B47	
A01_06_001	8 NUM	B19	B19	B19	A01_06_001
A01_06_002	8 NUM	B20	B20	B20	A01_06_002
A01_06_003	8 NUM	B21	B21	B21	A01_06_002
A01_06_004	8 NUM	B22	B22	B22	A01_06_004
A01_06_005	8 NUM	B24	B24	B24	A01_06_005
A01_06_006	8 NUM	B25	B25	B25	A01_06_006
A01_06_007	8 NUM	B26	B26	B26	A01_06_007
A01_06_008	8 NUM	B27	B27	B27	A01_06_008
A01_06_009	8 NUM	B28	B28	B28	A01_06_009
A01_06_010	8 NUM	B29	B29	B29	
A01_06_021	8 NUM				A01_06_021
A01_06_022	8 NUM				A01_06_022
A01_11_001	1 NUM	B31	B31	B31	
A01_12_001	8 NUM		B31a	B31a	A01_12_001
A01_07_001	8 NUM	B33	B33	B33	
A01_07_002	8 NUM	B34	B34	B34	
A01_07_003	8 NUM	B35	B35	B35	A01_07_003
A01_07_004	8 NUM				
A01_08_001	8 NUM	B64	B58	B58	A01_08_001
A01_08_002	8 NUM	B65	B59	B59	A01_08_002
A01_08_003	8 NUM				
A01_08_004	8 NUM				
A01_09_001	8 NUM	CEBELJE_DRUZINE	CEBELJE_DRUZINE	CEBELJE_DRUZINE	A01_09_001
A01_10_001	8 NUM	B66			
A01_10_002	8 NUM	B67			
A01_10_003	8 NUM				A01_10_003
A01_10_004	8 NUM				A01_10_004
A02_01_001	8 NUM	C7	C2	C2	A02_01_001
A02_01_002	8 NUM	C8	C3	C3	A02_01_002
A02_01_003	8 NUM	C9	C4	C4	A02_01_003
A02_01_004	8 NUM	C10	C5	C5	A02_01_004
A02_01_005	8 NUM	C11	C6	C6	A02_01_005
A02_01_006	8 NUM	C12	C7	C7	A02_01_006
A02_01_007	8 NUM	C13	C8	C8	A02_01_007
A02_01_008	8 NUM	C14	C9	C9	A02_01_008
A02_01_009	8 NUM	C15	C10	C10	A02_01_009
A02_01_010	8 NUM	C16	C11	C11	A02_01_010
A02_01_011	8 NUM	C17	C12	C12	A02_01_011
A02_01_012	8 NUM				A02_01_012
A02_01_013	8 NUM	C18	C13	C13	
A02_01_014	8 NUM	C19	C14	C14	
A02_01_015	8 NUM	C20	C15	C15	A02_01_015
A02_01_016	8 NUM	C21	C16	C16	A02_01_016
A02_01_017	8 NUM	C22	C17	C17	
A02_01_018	8 NUM				A02_01_018
A02_01_019	8 NUM	C23	C18	C18	
A02_01_020	8 NUM	C24	C19	C19	
A02_01_021	8 NUM	C25	C20	C20	
A02_01_022	8 NUM	C26	C21	C21	A02_01_022
A02_01_023	8 NUM	C27	C22	C22	A02_01_023
A02_01_024	8 NUM	C28	C23	C23	A02_01_024
A02_01_025	8 NUM	C29	C24	C24	A02_01_025
A02_01_026	8 NUM				A02_01_026
A02_01_027	8 NUM				A02_01_027
A02_01_028	8 NUM	C30	C25	C25	A02_01_028
A02_01_029	8 NUM	C31	C26	C26	A02_01_029

A02_01_030	8 NUM	C32	C27	C27	A02_01_030
A02_01_031	8 NUM	C33	C28	C28	A02_01_031
A02_01_032	8 NUM	C34	C29	C29	A02_01_032
A02_01_033	8 NUM	C35	C30	C30	A02_01_033
A02_01_034	8 NUM	C36	C31	C31	A02_01_034
A02_01_035	8 NUM	C37	C32	C32	A02_01_035
A02_01_036	8 NUM	C39	C33	C33	A02_01_036
A02_01_037	8 NUM	C40	C34	C34	A02_01_037
A02_01_038	8 NUM	C41	C35	C35	A02_01_038
A02_01_039	8 NUM	C42	C36	C36	A02_01_039
A02_01_040	8 NUM	C43	C37	C37	A02_01_040
A02_01_041	8 NUM				A02_01_041
A02_01_042	8 NUM				A02_01_042
A02_01_043	8 NUM	C44	C38	C38	
A02_01_044	8 NUM	C45			
A02_01_045	8 NUM	C46			
A02_01_046	8 NUM	C47	C41	C41	A02_01_046
A02_01_047	8 NUM	HMELJ	HMELJ	HMELJ	A02_01_047
A02_01_048	8 NUM	C65	C40	C40	A02_01_048
A02_01_049	8 NUM	C64	C39	C39	
A02_01_050	8 NUM				A02_01_050
A02_02_001	8 NUM	C49_1	C43_1	C43_1	A02_02_001
A02_03_001	8 NUM	C49_2	C43_2	C43_2	A02_03_001
A02_04_001	8 NUM				A02_04_001
A02_04_002	8 NUM				A02_04_002
A02_05_001	8 NUM	C49_3	C43_3	C43_3	
A02_02_002	8 NUM	C50_1	C44_1	C44_1	A02_02_002
A02_03_002	8 NUM	C50_2	C44_2	C44_2	A02_03_002
A02_01_051	8 NUM				A02_01_051
A02_02_003	8 NUM	C52_1	C46_1	C46_1	A02_02_003
A02_03_003	8 NUM	C52_2	C46_2	C46_2	A02_03_003
A02_04_003	8 NUM				A02_04_003
A02_04_004	8 NUM				A02_04_004
A02_05_002	8 NUM	C52_3	C46_3	C46_3	
A02_02_004	8 NUM	C53_1	C47_1	C47_1	A02_02_004
A02_03_004	8 NUM	C53_2	C47_2	C47_2	A02_03_004
A02_01_052	8 NUM				A02_01_052
A02_02_005	8 NUM	C55_1	C49_1	C49_1	A02_02_005
A02_03_005	8 NUM	C55_2	C49_2	C49_2	A02_03_005
A02_02_006	8 NUM	C56_1	C50_1	C50_1	A02_02_006
A02_03_006	8 NUM	C56_2	C50_2	C50_2	A02_03_006
A02_01_053	8 NUM				A02_01_053
A02_02_007	8 NUM	C58_1	C52_1	C52_1	A02_02_007
A02_03_007	8 NUM	C58_2	C52_2	C52_2	A02_03_007
A02_02_008	8 NUM	C59_1	C53_1	C53_1	A02_02_008
A02_03_008	8 NUM	C59_2	C53_2	C53_2	A02_03_008
A02_01_055	8 NUM	C61_1	C55_1	C55_1	A02_01_055
A02_03_009	8 NUM	C61_2	C55_2	C55_2	A02_03_009
A02_01_056	8 NUM				A02_01_056
A02_01_057	8 NUM				A02_01_057
A02_06_001	8 NUM	INT_SAD	INT_SAD	INT_SAD	
A02_06_002	8 NUM				A02_06_002
A02_06_003	8 NUM				A02_06_003
A02_06_004	8 NUM				A02_06_004
A02_06_005	8 NUM				A02_06_005
A02_06_006	8 NUM				A02_06_006
A02_06_007	8 NUM				A02_06_007
A02_07_001	8 NUM	C66	C57	C57	A02_07_001
A02_08_001	8 NUM	C67	C58	C58	A02_08_001
A02_10_001	8 NUM		C71	C71	A02_10_001
A02_10_002	8 NUM	C80			
A02_10_003	8 NUM	C81			
A02_10_004	8 NUM	C82			
A02_10_005	8 NUM	C83			
A02_12_001	8 NUM	C91			
A02_12_002	8 NUM	C92			
A02_12_003	8 NUM	C93			
A02_12_004	8 NUM		C74	C74	
A02_12_005	8 NUM	C94	C75	C75	
A02_12_006	8 NUM				A02_12_006
A02_12_007	8 NUM				A02_12_007
A02_12_008	8 NUM				A02_12_008

A02_12_009	8 NUM				A02_12_009
A02_12_010	8 NUM				A02_12_010
A02_09_001	8 NUM	C68	C59	C59	
A02_09_002	8 NUM	C69	C60	C60	
A02_09_003	8 NUM	C70	C61	C61	
A02_09_004	8 NUM	C71	C62	C62	
A02_09_005	8 NUM	C72	C63	C63	
A02_09_006	8 NUM	C73	C64	C64	
A02_09_007	8 NUM	C74	C65	C65	
A02_09_008	8 NUM	C75	C66	C66	
A02_09_009	8 NUM	C76	C67	C67	
A02_09_010	8 NUM	C77	C68	C68	
A02_09_011	8 NUM	C78	C69	C69	
A02_11_001	8 NUM		C72	C72	A02_11_001
A02_11_002	8 NUM	C85			
A02_11_003	8 NUM	C86			
A02_11_004	8 NUM	C87			
A02_11_005	8 NUM	C88			
A02_14_001	8 NUM				A02_14_001
A02_14_002	8 NUM	C96	C77	C77	A02_14_002
A02_14_003	8 NUM	C97	C78	C78	
A02_14_004	8 NUM	C98	C79	C79	
A02_14_005	8 NUM	C99	C80	C80	
A02_14_006	8 NUM	SKUPNI_PASNIK_AR	SKUPNI_PASNIK_AR	SKUPNI_PASNIK_AR	A02_14_006
A02_15_000	8 NUM	C100_DEC	C81	C81	
A02_15_001	8 NUM	C101_DEC	C82	C82	A02_15_001
A02_16_001	8 NUM			C83	A02_16_001
A02_16_002	8 NUM	C102	C83		A02_16_002
A02_16_003	8 NUM	C103	C84		A02_16_003
A02_17_000	8 NUM	C104_DEC	C85	C85	
A02_18_001	8 NUM	C105	C86	C86	A02_18_001
A02_19_001	8 NUM	C106			A02_19_001
A02_20_001	8 NUM	C107	C87	C87	A02_20_001
A02_21_000	8 NUM	C108_DEC	C88	C88	
A02_22_001	8 NUM	C109	C90	C91	A02_22_001
A02_23_001	8 NUM	C110	C91	C92	A02_23_001
A02_24_001	8 NUM	C1_DEC			
A02_25_001	8 NUM	C2_DEC			
A02_26_001	8 NUM	C3_DEC			
A03_01_001	8 NUM	C112	C93	C122	A03_01_001

Annex 3. Article; The Integration of Administrative Data for the Identification of the Ownership of Agricultural Land

Črt Šuštar[1]*, Mojca Eremita[2]

[1]Statistical Office of Republic of Slovenia

[2]Statistical Office of Republic of Slovenia

- **ABSTRACT**

Ownership structure of the agricultural land is one of the indicators for functional and justified orientation of agricultural policy-making. Agricultural holdings often rent agricultural land due to insufficient own land, which needs to be statistically measured. Despite adequate administrative registers in the country, this is not explicitly captured in a single data source. Therefore, it was formerly estimated using statistical questionnaires. In the scope of the 2020 Agriculture Census, administrative source data modelling process was developed that will substitute survey questions. Two geospatial administrative data sources (Integrated Administration and Control System – IACS and Land Cadastre) were combined based on their spatial relations. Results were exported in tabular format and combined with two tabular administrative registers (Register of Agricultural Holdings and Real Estate Register). Questions that arose due to joint and mixed ownership of agricultural units were addressed by testing different aggregation methods. The results were validated using historical data and the most comparable approach was chosen for future use. Project results provide quality data on agricultural land ownership – owned and rented land – for the whole population of agricultural holdings at the micro level. Data can be analysed separately for each land use and agricultural holding type. The developed process can be modified and applied elsewhere (both agricultural and other fields) to generate statistics.

Keywords: agricultural land ownership, administrative data sources, geospatial modelling, data integration, geospatial data sources

- **Abbreviations and acronyms**

AC - Agricultural census

AH - Agricultural holding

AP - Agricultural plot (land parcel as defined by IACS)

CAP - Common Agricultural Policy

EU - European Union

FSS - Farm Structure Survey

GDS - geospatial data sources

GEOSTAT - initiative to establish a data and production infrastructure for geospatial statistics

IACS - Integrated Administration and Control System

INSPIRE - Infrastructure for Spatial Information in Europe

LC - Land cadastre

LP - Land parcel (as defined by Land cadastre)

LPIS - Land parcel identification system (IACS database)

NSO - National statistics office

RAH - Register of Agricultural Holdings

RER - Real Estate Register

UAA - Utilised agricultural area

1. INTRODUCTION

For agriculture, land is one of the main inputs in the production process. Economically, it is the most efficient wealth-generating asset for agriculture holdings (AH) (Debertin, 2012). However, farmers often (at least partially) rent their utilised agricultural area (UAA). For the Common Agricultural Policy (CAP) as a European Union (EU) steering mechanism (CAP, 2019) this has various (more or less explicit) consequences.

To mention a few: 1) Rental rates are a function of expected market returns and expected associated direct payments (i.e. subsidies). As a result, land related subsidies increase rental rates (as well as land prices; Ciaian et al., 2010; Patton et al. 2008). 2) Occurrence of renting blurs the distinction on who received the CAP money in the end – farmers or landowners? Although farmers directly receive these payments, they may pass on a considerable share to landowners via increased farmland rental rates. The literature refers to this mechanism as 'incidence' (e.g., Kirwan, 2009; Breustedt & Habermann, 2011). 3) It affects AH size essentially affecting the farm structure.

Therefore, it is important to measure to what degree renting occurs in practice. In the case of Slovenia (EU), this is being monitored through the Agricultural Census (AC) and through Farm Structure Surveys (FSS). The AC is a statistical survey used to collect exhaustive information on all agricultural holdings (AH) above a certain threshold in the country, while the FSS is a sample survey on the structure of agricultural holdings. In the last decade, the AC was conducted once in 2010 and FSSs were conducted twice in 2013 and 2016 (Seljak & Krajnc, 2012). For 2020, a full capture AC is scheduled, for the first time mostly as an administrative census. The data will primarily be collected using existing administrative registers thus reducing respondent burden.

While many adequate data sources are available, the information needed is not always included in one single source. In such cases, data integration processes are utilised. While tabular data integration is traditionally exercised by national statistics offices (NSOs), not all data are stored in such a format. For registers with emphasised spatial distribution, geospatial data format is commonly used.

Integrating geospatial data sources (GDS) requires different approaches and methodology (compared to tabular data) and has not yet entered the mainstream of official statistics (Moström & Wardzińska, 2015). Unlike tabular data, geospatial data contain information on spatial distribution in graphical format. Thus, different sources can be integrated based on their spatial relationship using geospatial modelling to create new information from the existing (Pfeffermann et al., 2015). The use of GDS within an NSO can be broadly divided into two different categories (Moström, 2014):

- Production of geospatial statistics, where the geospatial statistics itself is released as the end-use product or at least forms an essential part of the result and
- Production of official statistics where geospatial information and/or geospatial processing is involved at some stage of the production chain (as an input data) but not essentially part of the disseminated result.

In recent years, two broader initiatives have influenced the area, namely INSPIRE and GEOSTAT. INSPIRE was an EU directive that standardised and made readily available GDS at the EU level, which made the use more efficient (Cetl et al., 2019; INSPIRE, 2019).

GEOSTAT is an ongoing initiative (in the form of projects) that was taken jointly by Eurostat and the NSOs to incorporate the production of geospatial statistics into the various phases of the Generic Statistical Business Process Model (GEOSTAT, 2019; Haldorson and Moström, 2019). The former made use of GDS more efficient and the later provided methodology and guidelines to integrate these data in official statistics production processes (Haldorson, 2012).

The presented work is a case of integrating geospatial and tabular administrative sources and a step towards a more complete exploitation of readily available data.

2. METHODOLOGY

- USED ADMINISTRATIVE SOURCES

- **Geospatial sources**

All used GDS were interoperable under the INSPIRE directive standard.

- *Land Parcel Identification System (LPIS)*

It is a polygon based ESRI shapefile managed by the Ministry of Agriculture, Forestry and Food (part of Integrated Administration and Control System – IACS database). Polygons represent (subsidized) agricultural plots (AP). The attribute part contains information on land use and identifier of subsidy receiver for each unit (LPIS, 2019). The source has a new version approximately every three months.

- *Land Cadastre (LC)*

It is a polygon based ESRI shapefile managed by the Surveying and Mapping Authority of the Republic of Slovenia (Ministry of the Environment and Spatial Planning). It is the official register with different layers available, depending on the purpose of use. In our case, we used a layer that continuously covers the area of the whole country where polygons represent informative borders of land parcels (LP). As such it is suitable for geospatial modelling. In Attributive table, LPs identifiers are included but not owner IDs (Ferlan & Vugrin, 2013; INSPIRE, 2019).

- **Tabular sources**

- *Register of Agricultural Holdings (RAH)*

It is an administrative source managed by the Ministry of Agriculture, Forestry and Food (MKGP, 2011). Information on AP subsidy receiver and its household members (15 or more years old) was used. It is continually updated.

- *Real Estate Register (RER).*

It is a multipurpose administrative data collection managed by the Surveying and Mapping Authority of the Republic of Slovenia – Ministry of the Environment and Spatial Planning that reflects the actual state of real estate in nature. A real estate can either be a parcel, a

parcel with a building, a building or a part of a building. It contains identifiers on real estate and owners. The data are updated weekly (Ferlan & Vugrin, 2013).

- PROCESS

The data integration process ran in two phases: geospatial modelling and tabular analysis (Appendix A). It was followed by data aggregation and data validation.

- **Geospatial modelling phase**

In the geospatial modelling phase layers of LC and LPIS were intersected in order to obtain all the LPs on each AP. The areas of AP and intersections were calculated. The resulting attribute table was exported in tabular format and processed in tabular phase.

- **Tabular phase**

The tabular analysis phase ran in two (2) steps; tabular integration of registers and aggregation.

1. In the first step RAH was used to identify AP user (subsidy receiver) and to each AH ID was assigned.
2. RER was used to identify LP owners. Records were multiplied in the case of co-ownership.
3. If an LP owner was a head of AH or an AH household member (15 or more years old, RAH data), AH ID was assigned to them. As not all LP owners are connected to AH, not every LP owner got an AH ID assigned.
4. For each record, LP owner AH ID and AP user AH ID were compared. If they were matching, LP (or part of a LP) that the record accounted for was counted as own, otherwise as rented.

- **Aggregation**

In the aggregation step, three sequential aggregations were made: 1) aggregation by LP (Real Estate Register), 2) aggregation by AP (IACS) and 3) aggregation by AH, using seven different aggregation methods (Table 1). Different aggregation possibilities arose due to joint LP ownership and multiple LPs (with different owners) on one AP.

Table 1 Aggregation methods[1]

Ownership share	**LP Minimal share**	**LP Major share**	**LP Real share**
AP Minimal share	Aggregation 1	Aggregation 4	*omitted (illogical)*
AP Major share	Aggregation 2	Aggregation 5	*omitted (illogical)*
AP Real share	Aggregation 3	Aggregation 6	Aggregation 7

[1] *Minimal share:* if at least one part of LP/AP was owned, it was counted as own. *Major share:* if at least one half of LP/AP was owned, it was counted as own. *Real share:* real share of LP/AP ownership was used for aggregation.

- **Validation**

Our aim was to match the model data as closely as possible to self-reported data. The results of different aggregation methods were validated through three indicators using historical data (AC 2010 – full survey, FSS 2013 and FSS 2016 – sample surveys) as ground truth (due to lack of more recent data). While there are no recent FSS data available, a model could theoretically be run on historical data to match the FSS time cross section. However, obtaining historical versions of administrative sources was not feasible in all cases.

Despite expected discrepancies, we assumed aggregation that provided results closest to the FSS data to be the most correct. It is not plausible that the renting structure during the gap period changed in such a way to make this assumption false. Three main indicators were taken into account (in all cases, pairs of each aggregation method and each FSS year were compared).

- *Indicator 1*

With $Indicator_1$, overall results were validated using differences in proportions of total rented UAA. For sample surveys 2013 and 2016 weighted area was used to make the proportion representative.

$$Proportion\ of\ rented\ UAA = \frac{\sum Rented\ UAA_{model/FSS}}{\sum All\ UAA_{model/FSS}}$$

Discrepancies between the proportions were calculated

$$Indicator_1 = Proportiopn\ of\ rented\ UAA_{model} - Proportion\ of\ rented\ UAA_{FSS}$$

- *Indicator 2*

With $Indicator_2$, differences at the AH level were accounted for. For each AH, only observations with data on both model and FSS were taken into account.

Two variations were calculated:

1) Sum of absolute differences of total UAA ($Indicator_{2totalUAA}$) does not depend on the model but estimates differences between LPIS and FSS data:

$$Indicator_{2\ total\ UAA} = \frac{\sum|Total\ UAA_{model} - Total\ UAA_{FSS}|}{UAA_{FSS}}$$

2) Sum of absolute differences of rented UAA ($Indicator_{2rentedUAA}$) to estimate the comparability of reported to model rented UAA.

$$Indicator_{2\ rented\ UAA} = \frac{\sum|Rented\ UAA_{model} - Rented\ UAA_{FSS}|}{UAA_{FSS}}$$

The absolute sum was chosen ahead of the sum of squares as we did not want to emphasize AH that actually changed renting during gap years (thus having a relatively big difference). No sample weights could be used (not to deform results at the AH level); however, the results depended on the total area (different for different FSS). Therefore, total sum was normalised dividing by FSS UAA.

- *Indicator 3*

With $Indicator_3$, occurrence frequency of farms with rented land was compared. Proportion of such farms was calculated:

$$Relative\ frequency\ of\ farms\ with\ rented\ UAA = \frac{\sum n_{Farms\ with\ rented\ UAA}}{\sum n_{All\ farms}}$$

For sample surveys 2013 and 2016 weighted frequencies were used in order to make the results representative:

$$Relative\ frequency\ of\ farms\ with\ rented\ UAA_{weighted} = \frac{\sum Weight_{Farms\ with\ rented\ UAA}}{\sum Weight_{All\ farms}}$$

Discrepancies between the relative frequencies were calculated:

$$Indicator_3 = Relative\ frequency\ of\ farms\ with\ rented\ UAA_{model} - Relative\ frequency\ of\ farms\ with\ rented\ UAA_{FSS}$$

3. RESULTS

- INDICATORS

- **Indicator 1**

Table 2: Indicator 1.

Model\Ground truth	AC 2010	FSS 2013	FSS 2016
Aggregation 1	0.063	0.043	0.040
Aggregation 2	0.141	0.121	0.118
Aggregation 3	0.166	0.146	0.143
Aggregation 4	0.084	0.064	0.061
Aggregation 5	0.162	0.142	0.139
Aggregation 6	0.186	0.166	0.163
Aggregation 7	0.198	0.178	0.175

Compared to all three FSS data, discrepancies between proportions of total rented UAA were the smallest in the case of aggregation 1 by a big margin (Table 2). According to aggregation 1, relative differences in rented UAA are 6%, 4% and 4% compared to AC 2010, FSS 2013 and FSS 2016, respectively.

- **Indicator 2**

Table 3: Indicator 2 total UAA. Total UAA did not depend on the aggregation method, therefore one result was compared.

Model\Ground truth	AC 2010	FSS 2013	FSS 2016
Aggregation	0.255	0.207	0.146

As expected, the more recent the ground truth data the smaller the average differences in total UAA on individual AHs (Table 3).

Table 4 Indicator 2 rented UAA

Model\Ground truth	AC 2010	FSS 2013	FSS 2016
Aggregation 1	0.199	0.227	0.207
Aggregation 2	0.215	0.229	0.206
Aggregation 3	0.228	0.232	0.212
Aggregation 4	0.204	0.224	0.204
Aggregation 5	0.224	0.234	0.214
Aggregation 6	0.237	0.237	0.219
Aggregation 7	0.246	0.242	0.224

Aggregation 1 proved to be closest to the 2010 AC data. For 2013 and 2016 FSS data aggregation 4 is ahead of aggregation 1 by a small margin (Table 4). According to aggregation 1, average differences in rented UAA on individual AHs are 20%, 23% and 21% compared to AC 2010, FSS 2013 and FSS 2016, respectively. According to aggregation 4, average differences in rented UAA on individual AHs are 20%, 22% and 20% compared to AC 2010, FSS 2013 and FSS 2016, respectively.

- **Indicator 3**

Table 5: Indicator 3

Model\Ground truth	AC 2010	FSS 2013	FSS 2016
Aggregation 1	0.145	0.078	0.078
Aggregation 2	0.222	0.156	0.156
Aggregation 3	0.684	0.617	0.617
Aggregation 4	0.167	0.100	0.101
Aggregation 5	0.242	0.175	0.175
Aggregation 6	0.684	0.617	0.617
Aggregation 7	0.687	0.621	0.621

Compared to all three ground truth data, discrepancies between relative frequencies of AH with rented land were smallest in the case of aggregation 1 by a big margin (Table 5). According to aggregation 1, the occurrence of AH with rented land changed by 14%, 8% and 8% compared to AC 2010, FSS 2013 and FSS 2016, respectively.

- Aggregation choice

In most cases, aggregation 1 proved to be closest to ground truth data by a big margin. An exception was Indicator 2 for FSS 2013 and 2016 but the margin was small. Therefore, aggregation 1 was chosen for future use. It implies that if at least one part of the plot is at least partially owned (or in co-ownership) by subsidy receiver (or one of their household members), the AP is considered owned and not rented.

4. DISCUSSION AND CONCLUSION

The used administrative data sources represent slightly different time cross sections. For development we chose versions that are less than a month to each other (June-July 2019). The choice of the versions in the production phase has to be done according to the survey methodology (reference date).

Model data are, compared to the ground truth data, of higher precision, allowing also land use differentiation. Accuracy was assessed through presented indicators. Compared to the FSS 2016 data (most recent), there is in total 4% more rented UAA ($Indicator_1$) and 8% more AHs that rented land ($Indicator_3$). All indicators reflect that more recent FSS data are closer to the model data. This indicates that the discrepancies are a consequence of changes in ownership and renting structure over time rather than a consequence of a model error. Apart from real changes in structure, there are several possible reasons for discrepancies. (1) AC and FSS had a threshold that was not accounted for in the model (full capture). (2) For the testing phase, only household members that were 15 or more years old were accounted. These data are readily available in RAH, while household members of all ages need to be identified from additional data sources. This is planned for the production phase. (3) In the case of deceased LP owners, the legal inheritance procedures are often lengthy. Therefore, the legal state (registers) and perceived ownership by the managers (self-reported data) may differ. (4) Additional reason for discrepancies is survey error in AC and FSS.

Using the procedure described above we managed to make a full capture of a state for a reference date. It was developed for the needs of the AC 2020 but the results can be reused in other areas of agricultural statistics.

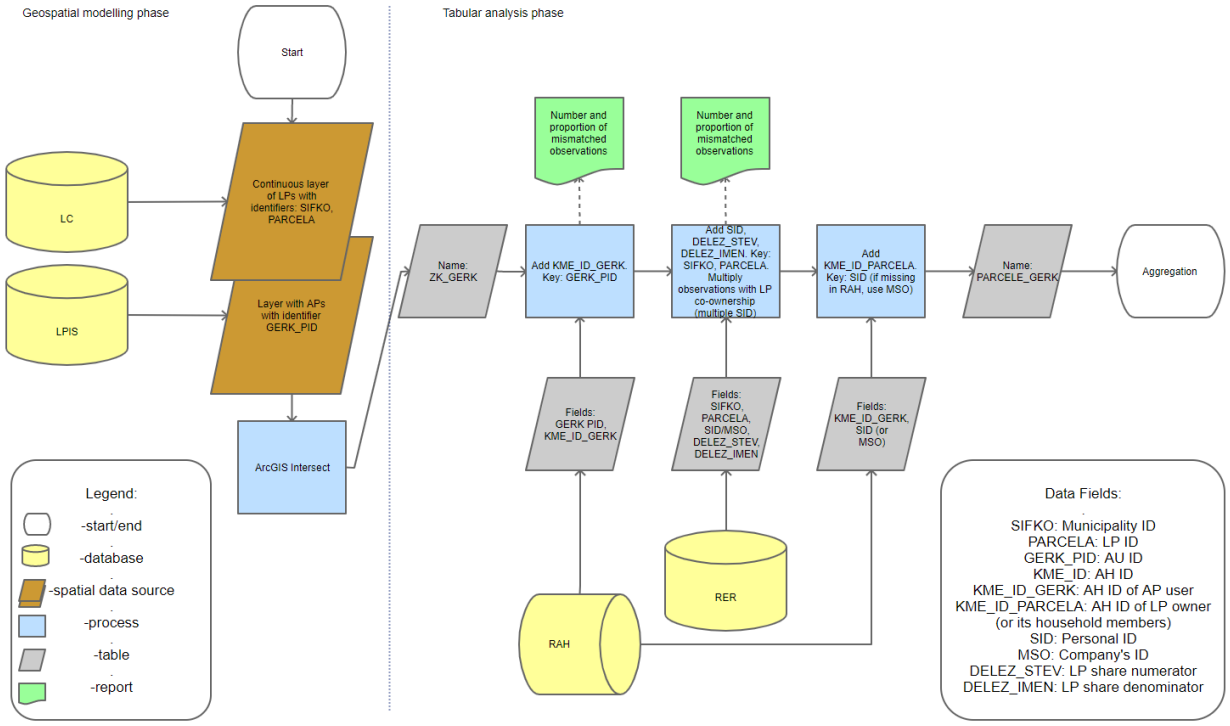
Most time and effort went into development and testing, while re-running the process is less work intensive. This enables us to make a full capture quickly and efficiently at any given time to obtain recent data without creating respondent burden.

Utilisation of administrative data ahead of surveys is what NSOs aim for; GDSs were, however, not commonly used before in agricultural statistics. Although extracting information from a single GDS can be done by tabular analysis, geospatial modelling requires specific software and knowledge. In order to fully utilise available data, it needs to become an integral part of the statistical data collection process.

- ## REFERENCES

1. Breustedt, G., & Habermann, H. (2011). The incidence of EU per-hectare payments on farmland rental rates: A spatial econometric analysis of German farm-level data. Journal of Agricultural Economics, 62(1), 225-243.
2. CAP, Common Agricultural Policy at a Glance. https://ec.europa.eu/info/food-farming-fisheries/key-policies/common-agricultural-policy/cap-glance_en visited Oct 2019
3. Cetl V., Tomas R., Kotsev A., de Lima V.N., Smith R.S., Jobst M. (2019) Establishing Common Ground Through INSPIRE: The Legally-Driven European Spatial Data Infrastructure. In: Döllner J., Jobst M., Schmitz P. (eds) Service-Oriented Mapping. Lecture Notes in Geoinformation and Cartography. Springer, Cham
4. Ciaian P., Kancs A., Swinnen J.F.M. (2010) EU Land Markets and the Common Agricultural Policy. CEPS Paperbacks. Available at SSRN: https://ssrn.com/abstract=1604452
5. Debertin, D. L. (2012) Agricultural Production Economics. Agricultural Economics Textbook Gallery. 1. https://uknowledge.uky.edu/agecon_textbooks/1
6. Ferlan M., Vugrin M. (2013) Javne nepremičninske evidence: Gradivo za strokovni izpit – geodetska stroka. Inženirska zbornica Slovenije, Ljubljana
7. GEOSTAT Initiative. https://ec.europa.eu/eurostat/web/gisco/gisco-activities/integrating-statistics-geospatial-information/geostat-initiative visited Oct 2019
8. Haldorson M. (2012) Implementing INSPIRE at Statistics Sweden and benefits from the INSPIRE data sharing, INSPIRE conference in Istanbul 2012
9. Haldorson M., Moström J. (2019) Developing a Statistical Geospatial Framework for the European Statistical System. In: Döllner J., Jobst M., Schmitz P. (eds) Service-Oriented Mapping. Lecture Notes in Geoinformation and Cartography. Springer, Cham
10. INSPIRE, Infrastructure for spatial information in Europe. http://inspire.ec.europa.eu/ visited Oct 2019
11. Kirwan, B.E. (2009). The Incidence of U.S. Agricultural Subsidies on Farmland Rental Rates. Journal of Political Economy 117: 138-164.
12. Krajnc A., Seljak R. (2012) Editing of multiple source data in the case of Slovenian agricultural census 2010. United Nations Economic Commission for Europe: Conference of European Statisticians
13. LPIS Technical description. https://www.sinergise.com/sites/all/themes/sinergise/images/LPIS-Technical_description-WEB.pdf visited Oct 2019
14. MKGP (2011) Aplikacija za vodenje Registra kmetijskih gospodarstev: Priloga 1 k Priročniku o vodenju registra kmetijskih gospodarstev.
15. Moström J., Wardzińska-Sharif A. (2015) Spatialising statistics in the ESS. Results from the 2015 GEOSTAT 2 survey on geocoding practices in European NSIs
16. Moström, J. (2014) More and better–Improvement of official statistics through the Swedish Geodata Cooperation. http://www.q2014.net/fileadmin/user_upload/More_and_better_-__Paper.docx visited Oct 2019
17. Patton, M., Kostov, P., McErlean, S., & Moss, J. (2008). Assessing the influence of direct payments on the rental value of agricultural land. Food Policy, 33(5), 397-405.
18. Pfeffermann, D., Eltinge, J. L., Brown, L. D., & Pfeffermann, D. (2015). Methodological issues and challenges in the production of official statistics: 24th annual Morris Hansen lecture. Journal of Survey Statistics and Methodology, 3(4), 425-483.

- # APPENDIX A: PROCESS WORKFLOW



Annex 4. Workshop "Modernisation of agricultural statistics"

Workshop regarding "Modernisation of agricultural statistics"; EU Grant number 2018.0218.

Dates: 8–9 October 2019

Location: Olsztyn, Poland

Participants from Slovenia: Statistical Office of the Republic of Slovenia – Agriculture, Forestry, Fishing and Hunting Statistics Section

1. Summary of the workshop

- Eurostat:
  - There is a tendency to validate data on the source and implement a centralized data warehouse.
  - There is a strong need to solve open issues on official statistics data access to private big data sources.
  - Geodata and georeferenced statistics are becoming increasingly emphasized.
  - There is a need for guidelines for multiple source data validation.
- Geodata and satellite imagery:
  - Poland: The Register of Agricultural Holdings is being extended using geospatial data (land cadaster, location of agricultural holdings, areas with restricted land use, etc.) and the Population Register. The use of satellite imagery is under development in the scope of pilot projects (in cooperation with IGIK, ESA and Space Research Centre):
    - Early crop estimation - Using NDVI (vegetation index) based on Sentinel data (10 days revisit time). Currently it is a parallel system validated by expert estimates.
    - Crop recognition using drones: there are limitations regarding land size capture capabilities (time and range of flight).
    - Crop recognition using Sentinel data (1&2) combined with administrative data. Promising results for land parcel size >0.5ha and wider than 20m (90% accuracy). For smaller land parcel size acquisition higher resolution imagery is necessary. Higher resolution also significantly increases time and computation capacity needs; therefore, a combined approach is used. Images captured by the end of July are used (later images do not add significantly to the accuracy), processing time is currently 5-6 weeks. Methodology will be proposed to Eurostat for validation and confirmation.
    - Pilot project with Estonia for treatment recognition (irrigation, mowing, etc.) is in its initial stage.
  - Germany: Pilot project SatAgrarStat for yield statistics with the aid of Sentinel 2 data for three main crops. For validation, crops are being weighted and reported. Models achieve sufficient accuracy but need to be improved for the cases of extreme weather events (draughts, floods, hail, etc.).
  - Finland:
    - Analysis of crop rotation using LPIS (Land parcel identification system - IACS database). Where the same crop is cultivated for three years - no crop rotation. For mixed crop areas, the average of minimal and maximal overlap is counted.
    - Location of agricultural holdings: calculated using LPIS plots. If the agricultural plots are more than 30km from one another,

separate clusters are formed. As location, the biggest plot of the biggest cluster is taken.
    - Latvia: pilot projects are under development for Sentinel data use: control the administrative data (LIPIS) and identification of agricultural use not in registers.
- Other:
    - Greece: GIS applications are used for coordinates of agricultural holdings. For census (CAPI), tablets are used.
    - Finland: for web questionnaires, identification system of the Finnish central bank was implemented (usernames and passwords are no longer in use). They see a big potential in farm management software data (national legislation already enables them access).
    - Poland: validation rules are a part of the metadata system.
    - Extending administrative registers (e.g. Register of Agricultural Holdings) using other data sources (insurance data, taxes, cadaster, and satellite data). E.g., market gardeners that are not agricultural holdings are obtained from insurance registers in Germany, and from tax registers in Poland.
- Discussion:
    - In most countries, Statistical Registers of Agricultural Holdings are updated only through administrative sources. There was a consensus that the occurrence of significant farms not included in the register is negligible.
    - Administrative data and statistical needs are often discrepant, which is solved by coordination as well as modelling.
    - Many countries have issues with many farms on the same address. That is not the case in Slovenia as the statistical definition in this case is in accordance with the Ministry.
    - With the new agricultural policy it is possible that the current administrative data collections would significantly shrink.
    - The need for better knowledge and experience exchange for the case of satellite data use in statistics was communicated. There is also a need to implement official methodologies for that area. Grants are helpful for those aspects but are financially insufficient.
    - Enterprises working in big data for agriculture (precision farming, satellite data, etc.) often have dubious methodologies, so caution is needed when using those data.

2. Agenda

8 October

7:00 – 9:00 Registration

9:00 – 13:00 Opening session

9:00 – 10:30

- Introduction/welcome - Artur Łączyński, Director of Agriculture Department Statistics Poland, Marek Morze, Statistics Poland, Director of the Statistical Office in Olsztyn
- “The role of Eurostat in modernizing agricultural statistics” - Marjo Kasanko, Deputy Head of the Unit Agriculture and Fisheries, Eurostat
- “Agricultural census 2020 in Greece” – Lemonia Dionysopoulou, Head of the Agriculture, Livestock, Fishery and Environmental Statistics Division, ELSTAT

- Discussion

10:30 – 11:00 Coffee break

11:00 – 13:00

- "Modernisation of agricultural statistics " – Anna-Kaisa Jaakkonen, Senior Statistician, and Arja Anttila, Senior Specialist, Luke Finland
- "Modernisation of agricultural statistics" – Kaspars Misans, Vice-president, Central Statistical Bureau of Latvia
- "Modernisation of agricultural statistics" – Artur Łączyński, Director of the Agriculture Department, Statistics Poland
- Discussion
- Summary/conclusions – moderator and participants

13:00 – 14:00 Lunch break

14:00 - 16:30 Session 1. Modern methods of data collection

14:00 – 15:00

- Introduction – moderator Pasi Mattila, Senior Scientist, Luke Finland
- "Reduction the burden of respondents and improvement of statistics quality by effective use of administrative data sources" – Anita Raubena, Deputy Director of the Agricultural and Environmental Statistics Department, Central Statistical Bureau of Latvia
- "Farm planning software and internet services as a potential source of field-parcel-specific data for agricultural statistics" – Pasi Mattila, Senior Scientist, Luke Finland

15:00 – 15:30 Coffee break

15:30 – 16:30

- Discussion
- Summary/conclusions – moderator and participants

9 October

9:00 – 10:30 Session 2. Geospatial and satellite data

- Introduction – moderator Artur Łączyński, Director of the Agriculture Department, Statistics Poland
- Use of geospatial field parcel data of IACS in the investigation of crop rotation" – Pasi Mattila, Senior Scientist, Luke Finland
- "Possibilities of using satellite data for agricultural statistics' needs" – Tomasz Milewski, Expert, Statistics Poland
- "Project on satellite use for harvest estimation in Germany'' – Ute Walsemann, Germany
- Discussion
- Summary/conclusions – moderator and participants

10:30 – 11:00 Coffee break

11:00 – 12:30 Session 3. Methodological issues

- Introduction – moderator Barbara Domaszewicz, Deputy Director of the Agriculture Department, Statistics Poland
- "The concept of data validation mechanism developed/within the grant project: Implementation of business architecture for ESS Validation" - Dariusz Miziołek, Expert, Statistics Poland; Marek Panfiłow, Expert, Statistics Poland (Statistical Office in Olsztyn)
- "Merging farm register data and administrative data" – Anna-Kaisa Jaakkonen, Senior Statistician, Luke Finland
- Discussion
- Summary/conclusions – moderator and participants

12:30 – 13:30 Lunch break

13:30 – 14:30 Session 4. Modernization of the statistical farm register (SFR)

- Introduction – moderator Kaspars Misans, Vice-president, Central Statistical Bureau of Latvia
- "Statistical farm register system"– Anneli Partala, Senior Statistician, Luke Finland
- Discussion
- Summary/conclusions – moderator and participants

14:30 – Closing session – summary