



Progress Report

(for period: 09/2021 - 08/2022)

Action title: Administrative data sources for agricultural statistics

Modernisation of agricultural statistics

Grant Action Number	101036501
Grant Action Acronym	2020-CZ-AGRI
Starting date	01/09/2021
End date	31/08/2024
Duration in months	36
Call (part) identifier	ESTAT-2020-PA8-E-AGRI
Topic	B5476-2020-AGRI

Contact person: Radek Matějka (radek.matejka@czso.cz)

31 August 2022

This document has been produced with the financial assistance of the European Commission. The views expressed herein are those of the author and can therefore in no way be taken to reflect the official opinion of the European Commission.

Table of contents

Table of contents	2
Introduction	3
General comment on the progress of the Grant Action	3
Available administrative data sources (ADS)	3
Tasks covered by the interim Progress Report	
Description of work done, interim results and conclusions	
- Animal production statistics	
- Crop production statistics	
Further work envisaged	

Introduction

The Czech Statistical Office (CZSO) is implementing a Grant Action (further called also a project) aimed at modernizing the system of agricultural statistics by establishing and implementing a sustainable system for the regular use of data from administrative data sources (ADS).

The primary aim of the project is to reduce the response burden of agricultural units and rationalize available resources while maintaining the high quality of agricultural statistics. Another equally important goal is to establish new or deepen existing cooperation between CZSO and other state institutions that own and administer the administrative data sources, which are to replace existing data collection systems - statistical sample surveys.

Project activities are divided into 3 Objectives:

Objective A (Tasks A.1 – A.7): Mapping and analyses of ADS

Objective B (Tasks B.1 - B.6): Identification and specification of requirements for a sustainable system for the use of administrative data

Objective C (**Tasks** C.1 - C.4): Establishing and implementation of a sustainable system for export and integration of administrative data

Period covered by the Progress Report: 01/09/2021 – 31/08/2022

General comment on the progress of the Grant Action

- The Grant Agreement for the project was signed at the beginning of September 2021.
- End of September 2021 the Czech Statistical Office received a pre-financing from Eurostat to cover eligible costs of the project according to agreed budget.
- The kick-off coordination meeting of the Interim working group, where only CZSO experts can participate, was organized 22 September 2021. The Grant Action Manager chaired the meeting. At the first coordination meeting the technical details (organization of work, timetable, planned meetings, coordination between experts and distribution of roles, involvement of experts in each task, etc.) administrative issues (worksheets, budget plans for the period of Sept Dec 2021) were provided. Experts presented the summary of the recent situation in their topics (animal statistics, crop statistics, a Farm Register). The Grant Action Manager informed about planned interinstitutional meeting of External working group (external experts from state institutions owning the administrative data sources).
- Members of the realization team are statisticians experts on animal production and crop production statistics, on structural agricultural statistics and on agricultural registers.
- Coordination meetings of the Internal working group are organized with the frequency of 2 to 4 weeks, as needed following the progress of the project.
- The coordination meetings have the following agenda: a summary of work results of the project in each topic, encountered problems and solutions to solve them, planned actions and tasks for the nearest weeks.
- 14 coordination meetings took place during the period September 2021 August 2022.

Available administrative data sources (ADS)

Since the beginning of the Grant Action, the activities focused on the mapping and analyses of available administrative data sources:

1. Agricultural register LPIS (Land Parcel Identification System, further also called land register)

LPIS is a geographic information system (GIS) established on the basis of Act No. 252/1997 Coll., on agriculture, in 2004.

LPIS is primarily intended for keeping records on the use of agricultural land in the Czech Republic and for verifying data in applications for agricultural land-linked subsidies, regardless of whether the subsidies are financed from EU resources or national subsidy programs. Farmer's blocks (DPB) is a basic entity of the LPIS that represents a continuous area of agricultural land. Each DPB is managed by one farmer.

LPIS also serves, for example, as a basis for keeping records on the use of fertilizers, pastures, plant protection products, for determining management restrictions due to the nitrate directive, erosion risks, etc. LPIS is also used for the localization of outbreaks of animal diseases or for the monitoring of harmful organisms occurrence.

The Ministry of Agriculture (MoA) is the administrator of LPIS at the hardware and software level and the State Agricultural Intervention Fund (SAIF) is responsible for updating the land register from submitted subsidy applications.

2. IS SAIF database, the Information System of SAIF – a part of IACS (Integrated Administration and Control System)

The State Agricultural Intervention Fund is an accredited paying agency established and governed by Act No. 256/2000 Coll., on the State Agricultural Intervention Fund. SAIF is an intermediary of the financial support for agricultural holdings from the European Union and national sources. It keeps and administers a specialized information system (IS SAIF) on subsidy applications and subsidy recipients. The official name of the subsidy application form is the Single Application.

3. IAR register (Integrated Animal Register, further also called animal register) - a part of IACS

Integrated Animal Register is used to register farm animals in compliance with the legislative requirements on labeling animals, their relocation and declaring the origin of animal products, given by European and national legal acts:

- Regulation (EC) No 1760/2000 of the European Parliament and of the Council of 17 July 2000 establishing a system for the identification and registration of bovine animals and regarding the labelling of beef and beef products and repealing Council Regulation (EC) No 820/97
- Regulation (EU) 2016/429 of the European Parliament and of the Council of 9 March 2016 on transmissible animal diseases and amending and repealing certain acts in the area of animal health
- Commission Delegated Regulation (EU) 2019/2035 of 28 June 2019 supplementing Regulation (EU) 2016/429 of the European Parliament and of the Council as regards rules for establishments keeping terrestrial animals and hatcheries, and the traceability of certain kept terrestrial animals and hatching eggs
- Act No. 154/2000 Coll., on breeding and registration of animals and on the amendment of some related laws (Breeding Act)

• Decree No. 136/2004 Coll., establishing the details of labelling animals and their registration, registration of farms and persons as established by the Breeding Act

There are two primary purposes of the IAR. The register is used, firstly, for the registration of animals due to veterinary requirements on the prevention and suppression of the occurrence of infectious diseases of animals and, secondly, for the creation of an effective system for the traceability of products of animal origin. The IAR is also used to verify data in applications for animal-linked subsidies, regardless of whether the subsidies are financed from EU resources or national subsidy programs.

The IAR is administered and owned by the Ministry of Agriculture (MoA) and the Czech Moravian Breeders'Corporation is responsible for maintaining and updating the animal register.

Tasks covered by the interim Progress Report

Task No.	Task name	Description	Animal statistics	Crop statistics
A	Analysis	Analysis of administrative data sources IACS: for crop production statistics: LPIS (Land Parcel Identification System) IS SAIF (information system of the State Agricultural Intervention Fund) for animal production statistics: Identification and registration system: IAR (Integrated Animal Register)	V	Ø
A.1	Analysis of ADS technical elements • Concepts, definitions of variables, code list items. • Detection of main differences between the current legislation (Reg. (EC) 543/2009 (crop statistics), Reg. (EC) 1165/2009 (animal statistics), SAIO Regulation/SAIO Implementing acts) and ADS. • Assessment of differences, the definition of measures to eliminate or manage detected differences. • Estimating the possible limitations and negative impacts of using data from administrative sources. • Definition of measures to avoid or reduce negative impacts on data quality.		V	V
A.2	Analysis of user needs Analysis and evaluation of user needs (user satisfaction surveys).		V	V
A.3	ADS item/variable analysis	/variable extractable from ADS and those missing in ADS.		V
A.4	ADS Coverage	 Description of the ADS coverage degree at the level of individual statistical units (holdings/farms) and UAA (utilized agricultural area). Description of the minimum entry thresholds for ADS. Impact evaluation of different minimum entry thresholds for ADS and for crop production statistics. 	V	V

A.5	Updating of ADS	 Description and frequency of the ADS updating process. The timeliness of recording changes to ADS and their impact on data quality. For crop production statistics: Assessment of data processing and publishing systems (with or without regular data revisions). For animal statistics: Setting of a sufficient degree of data updating for provisional and/or final data, the alignment of reference date with time limits as set out in the legislation in force. Determining the optimal date for data extraction from ADS to match data quality requirements and deadlines laid down by the legislation in force. 	V	Ø
A.6	Data analysis	 Analysis of data available in ADS: Data comparison at the level of individual statistical units (holdings/farms). Comparison with data available in sample sets for CZSO statistical surveys and in the Farm Register kept by the CZSO. 		V
A.7	Study visits	Exchange of experience and good practice in the analysis of ADS in more experienced Member States.		×
В	Requirement formulation			✓ ongoing
B.1	Requirements for external working group			☑ ongoing
B.2	Output table preparation	 Defining the structure of output tables according to the requirements of national data users' needs and of the legislation in force. Mapping of the CZSO and ADS code list items. Setting/programming of aggregate items in output tables. Basic checks setting/programming. Prevention of duplicated records. Basic data checks between LPIS (main area) and IS SAIF (individual crop area). 	☑ ongoing	☑ ongoing

		 Data cross-validation against data recorded in the previous period. Setting rules for the confidential data protection. 		
B.3	Preparation of adjustments to the Statistical Metainformation System	• Preparation of adjustments to be implemented in the CZSO Statistical Metainformation System (SMS), related to the use of ADS (code lists, definitions, data processing, output tables, preparation of the system for the data use from ADS, modification of the statistical survey system).	☑ ongoing	☑ ongoing
B.4	Preparation of the solution for missing items	the solution for Additional sample survey - elaboration of methodology.		☑ ongoing
B.5	Methodology for updating the Farm Register Elaboration of the methodology for updating the Farm Register (kept by the CZSO) with data exported from ADS.		☑ ongoing	☑ ongoing
B.6	Update of the methodology for linked statistical surveys	Modification and extension of the methodology for linked statistical surveys in the crop production statistics (harvest estimate surveys, final harvest survey) and animal production statistics (surveys on cattle production, pig production and poultry production).	☑ ongoing	☑ ongoing

Description of work done, interim results and conclusions

- Animal production statistics

A.1 Analysis of ADS technical elements

The aim of this phase was to determine the transferability of data between the IAR and items in existing statistical surveys (i.e. the current source) and items according to the new SAIO regulation.

- the IAR stores information on individual animals (individual records) for cattle, sheep and goats, but on individual holdings/farms in case of pigs
- the required animal categories (items) are not directly available in the animal register, but can be determined in the following ways:
 - a) by combining various attributes stored in the animal register, e.g. the animal's date of birth, sex, delivery record, animal breeding system
 - b) by deriving from the categories determined under point (a)
 - c) by modeling, based on additional information, or
 - d) in some cases, however, it is not possible to determine the animal category in such a way to fit to the definition in the legislation
- for cattle, sheep and goats, data sources differ in the used observation unit. In the sample statistical survey, the observation unit is the holding/farm, but in the animal register, the observation unit is the animal.

The observation unit for pigs and poultry is the same in both sources (it is a holding/farm).

Factors limiting the utilization of the IAR for the distinction of animal categories:

The availability of information about animal categories in the IAR is affected by the purpose for which the animal register was established and is maintained. It means that the attributes necessary for distinguishing animal categories might not be included.

The current source (a sample statistical survey) ensures that the animal categories are defined according to the valid legislation.

Bovines

- Required items can be filled from the IAR.
- Distinguishing attribute used for dividing cows into *dairy cows* and *other cows* does not comply with the legislative requirements set for animal statistics. Dairy cows can be identified by the existence of a milk collection room on the holding/farm. Then the item *dairy cows* is deemed to include only cows milked for the purpose of placing milk and milk products on the market and not cows milked for own milk consumption; those cows are considered as other cows.

Pigs

- Only total number of pigs is available in the IAR.
- Pig categories according to live weight cannot be downloaded directly nor derived from IAR.
- No other indicator was identified to be used for modelling required categories while maintaining sufficient quality of estimates.

Sheep, goats

- Although there are individual records of sheep and goats in the IAR, the categories, as required by the statistical legislation, cannot be derived from the IAR.
- An attribute to be used for defining the *mated status* for ewe-lambs put to the ram and goats mated for the first time is not available.
- Modeling was carried out based on various parameters: (a) proportion of breeding ewes/goats, (b) a combination of registered delivery and the age of ewes/goats, (c) a usual number of ewe-lambs that have already lambed.
- Distinguishing attribute used for dividing sheep into *milk breeding females* and *non milk breeding females* does not comply with the legislative requirements set for animal statistics. Milk breeding females can be identified by the existence of a milk collection room on the holding/farm. Then the item *milk breeding females* is deemed to include only sheep milked for the purpose of placing milk and milk products on the market and not sheep milked for own milk consumption; those sheep are considered as non milk breeding females.

Other limiting factors of the IAR:

An animal keeper is the person responsible for entering data in the animal register. Mandatory and non-mandatory data recorded in the IAR are checked against the control links between data. In case of violation, the animal keeper has to correct the records. Control links are defined so that to ensure that the data entering the animal register corresponds with the IAR purpose.

Limiting factors identified in the data analysis:

- a) Arising from the discrepancy in mandatory attributes:
 - inactive animal keepers and their animals might be registered in the IAR
 - no-longer-alive animals might be registered in the IAR
 - date of birth of the animal might be missing in the IAR

Removal of these limiting factors will be required from the MoA.

- b) Arising from the discrepancy in non-mandatory attributes:
 - non-mandatory data related to attributes used for determining animal categories are not recorded correctly (e.g. data on delivery is missing)

The MoA will be asked to minimize these substantial limits.

A.2 Analysis of user needs

A meeting was held with data users (e.g. the economic accounts for agriculture, the regional yearbooks, a Report on Czech Agriculture of MoA, etc.), where the new time schedule of statistical data availability after the change in data source was discussed.

In the current system, the November/December data of year N published on 15 February of year N+1 are considered as definitive. When data is derived from the IAR, however, the data on 15 February N+1 will only be preliminary but definitive not earlier than on 30 April N+1. This change negatively affects the availability of final data for internal data users used to have them on disposal already in February. The requested data will be replaced with estimates or the publication date will be adequately shifted.

Additional information – see A.2 in Crop production statistics part.

A.3 Item/variable analysis

Availability of animal species and categories in the IAR

Table 1.

A2000	L	ive b	povine animals	Available
A2010		Boy	vine animals less than 1 year old	Available
A2010B		В	Sovine animals, less than 1 year old, for slaughter	Modelling
A2010C		В	Sovine animals, less than 1 year old, not for slaughter	To be derived
A2110C			Bovine male calves, less than 1 year old, not for slaughter	To be derived
A2210C			Female calves, less than 1 year old, not for slaughter	To be derived
A2020		Boy	vine animals 1 to less than 2 years old	Available
A2120		N	Male bovine animals, 1 to less than 2 years old	Available
A2220		Н	Ieifers, 1 to less than 2 years old	Available
A2220B			Heifers, 1 to less than 2 years old, for slaughter	Modelling
A2220C			Heifers, 1 to less than 2 years old, not for slaughter	To be derived
A2030		Boy	rine animals, 2 years old or over	Available
A2130		N	Male bovine animals, 2 years old or over	Available
A2230_2300	Female bovine, 2 years and over (including all cows)		emale bovine, 2 years and over (including all cows)	Available
A2230			Heifers, 2 years old or over	Available
A2230B			Heifers, 2 years old or over, for slaughter	Modelling
A2230C			Heifers, 2 years old or over, not for slaughter	To be derived
A2300			Cows	Available
A2300F			Dairy cows	Available, but the category definition differs from statistical legislation (dairy cows = milk collection room on farm)
A2300G			Non-dairy cows	Available, but the category definition differs from statistical legislation (other cows = farm without a milk collection room)
A2400		of v	vhich buffaloes	Non existing
A2410		В	reeding female buffaloes	Non existing
A2420		О	Other buffaloes	Non existing

A3100	Li	ve s	wine, c	Available		
-		Pig	gs of <	50 kg	Not available	
A3110			Piglet	s of <20 kg	Not available	
A3131			Other	pigs of <50 kg	Not available	
A3132		Fa	ttening	pigs, including cull boars and cull sows, of 50 kg and more	Not available	
A3132X			Fatten	ning pigs, from 50 kg to <80 kg	Not available	
A3132Y			Fatten	ning pigs, from 80 kg to <110 kg	Not available	
A3132Z			Fatten	ning pigs, of 110 kg or over	Not available	
A3120_3133		Br	reeding pigs, of 50 kg and more		Not available	
A3133			Breed	ling boars	Not available	
A3120			Breed	ling sows	Available, but the definition differs – gilts are not included	
A3120K			Co	overed sows	Not available	
A3120KA				Sows covered for 1st time	Not available	
A3120L			So	ows not covered	Not available	
A3120LA				Gilts not yet covered	Not available	

A4000	Li	ve s	heep and goats	Available
A4100		Liv	ve sheep	Available
A4110K			Breeding female sheep	Not available/modelling
A4110KC			Milk breeding female sheep	Not available/different definition
A4110KD			Not available/definition differs	Not available/different definition
A4120			Other sheep	Not available/modelling
A4200	Live goats		ve goats	Available
A4210K		Goats breeding females Not available/mo		Not available/modelling
A4220			Other goats	Not available/modelling

A.4 The IAR coverage

Bovines, sheep, goats

- Without threshold values.
- Individual registration of individual animals.

Pigs

- Without threshold values.
- Available data: number of pigs, of which sows.
- Data for a holding/farm, not for individual animals.

An administrative source without thresholds ensures a full update of the Farm Register (containing the frame population); however, a suitable identifier to link the two registers (IAR and the Farm Register) will need to be found and the solution of how to deal with the existence of inactive farms in both registers.

A.5 Updating of the IAR

The keepers report data to the IAR according to a prescribed time schedule with the deadlines defined in Decree 136/2004 Coll. as follows:

For cattle, sheep and goats

- Changes in number on animals (death, loss, culling, relocation) to be reported within 7 days of the occurrence of the event,
- a birth of the animal to be reported within 7 days of identification,
- identification of calves to be reported within 20 days of birth,
- identification of lambs and kids to be reported within 6 months of birth or when moved from the holding/farm.

For pigs

• Number of pigs as of the last day of the month.

This time schedule for updating records will affect the completeness of the statistical data, because the quality of downloaded data can differ depending on the time lag between the reference date and the extraction date. Optimization of the timeline "reference date - extraction date - publication date" will be linked to the reference date 31 December 2022.

A.6 Data analysis

Data from the IAR were downloaded within 1 to 2 months after the reference date of 1 April 2022 and subjected to a comparative analysis with the results of the CZSO's sample statistical survey *Livestock Survey*.

Extracted data contained attributes agreed between the CZSO and the MoA so that it was possible to complete as many of the required categories as possible (see part A3).

The structure of the data record contained:

- Identification of the unit.
- animal identification and information about the animal (ID, date of birth, sex, breed, delivery),
- identification and information about the holding/farm (ID of the holding/farm, cadastral territory, organic farming regime),
- other attributes (milk collection room on the farm)

Bovines

Results of the comparative analysis are shown in the following table.

Table 2.

	IAR	CZSO	Di	fference
Bovine categories	as at 1 April 2022	as at 1 April 2022	heads	% (100 = CZSO)
Live bovine animals	1 448 074	1 421 254	26 820	1,9
Aged less than 1 year	439 525	429 534	9 991	2,3
For slaughter	4 634	655	3 979	607,5
Not for slaughter	434 891	428 879	6 012	1,4
Male calves	180 963	176 653	4 310	2,4
Female calves	253 928	252 226	1 702	0,7
Aged 1 to less than 2 years	318 940	314 905	4 035	1,3
Males	107 739	104 912	2 827	2,7
Heifers	211 201	209 993	1 208	0,6
For slaughter	18 247	9 324	8 923	95,7
Not for slaughter	192 954	200 669	-7 715	-3,8
Aged 2 years or over	689 609	676 815	12 794	1,9
Males	24569	21079	3 490	16,6
Females	665 040	655 736	9 304	1,4
Heifers	77 137	67 877	9 260	13,6
For slaughter	6 665	2 311	4 354	188,4
Not for slaughter	70 472	65 566	4 906	7,5
Cows	587 903	587 859	44	0,0
Dairy cows	323 683	358 255	-34 572	-9,7
Non-dairy cows	264 220	229 604	34 616	15,1
of which buffaloes	0	0		
Breeding female buffaloes	0	0		
Other buffaloes	0	0		
of which unclassified	3 624			

Commentary on the table 2.:

- The biggest differences between compared sources were recorded in the *for slaughter* categories. In the current source, which is the CZSO sample survey, the determination of the category *for slaughter* is the result of the respondent's statement that he keeps animals that will be slaughtered in the relevant age category. However, this statement is largely dependent on the intention of the keeper and he may not know for sure on the reference day whether the animal will actually be slaughtered (e.g. heifer for slaughter) or will be included in the breeding (e.g. heifer for breeding). This approach significantly underestimates the number of animals intended for slaughter in the relevant age category.
- The animal register does not contain attributes allowing determining whether the animals are intended *for slaughter* or not.
- The existence of other sources of data on animals that have been slaughtered (slaughtering statistics and foreign trade in animals intended for slaughter) gives the possibility of modeling. The CZSO created a methodology for modeling the number of animals for slaughter based on the relationship between the gross domestic production of animals for slaughter in the previous (model) period and the circularity coefficient of the category. Numbers of animals *for slaughter* are expressed as average numbers over the model period.
- The data obtained from the modeling should better correspond to the number of animals that will be subsequently slaughtered than the number estimated by the respondent. In the event that modeling is used to generate the aforementioned statistics, they will be marked as breaks in the time series.
- Significant differences were noted in categories of animals over 2 years old, except for cows.
- This finding is closely related to the fact that over-aged animals, probably already nolonger-alive, or in inactive farms, are registered in the animal register.
- CZSO will negotiate with the MoA on the request for the maximum elimination of such animals from the animal register.
- On the contrary, minimal differences were observed between the total number of cows reported by the respondents and the total number of cows in the cow register.
- The animal register could be considered a very high-quality source of data on the total number of cows, however, the mentioned sources differ in determining the livestock yield (dairy x other). In the animal register, cows can be divided into dairy cows and other cows using the attribute *existence of a milk collection room* in the farmer's holding, proving that the farmer supplies milk or milk products to the market. Cows milked for own milk consumption are not recorded.
- CZSO will discuss with the MoA the possibilities of a more detailed distinction between dairy cows and other cows to better correspond to the definitions.
- The animal register contains a certain number of animals that cannot be classified in any
 of the categories because the records of non-mandatory but also mandatory attributes
 are missing. The CZSO will ask the MoA to focus on increasing the completeness and
 quality of the data kept in the animal register.

Sheep and goats

The category *breeding female sheep/goats* cannot be directly downloaded from the animal register because there is no appropriate attribute that would allow the category *breeding females*

put to the ram for the first time to be specified as a part of the category breeding female sheep/goats.

For the analysis below, the following coefficients were defined:

- for modeling the category of breeding females, which is the share of breeding females from the total number of sheep/goats in IFS 2020: 69.540% for sheep and 62.411% for goats,
- for sheep only: for modeling the category of milk breeding sheep, which is the share of females in farms with a milk collection room from the total number of females in the animal register: 1.822%.

Results of the comparative analyses are shown in the following tables.

Table 3.

		IAR	CZSO	Diffe	rence
Sh	eep categories	as at 1 April 2022	as at 1 April 2022	heads	% (100 = CZSO)
Liv	ve sheep	247 441	174 196	73 245	+42,0
	Breeding female sheep	172 070	112 982	59 088	+52,3
	Milk breeding female sheep	3 135	5 100	-1 965	-38,5
	Non milk breeding female sheep	168 935	107 882	61 053	+56,6
	Other sheep	75 371	61 214	14 157	+23,1

Table 4.

	IAR	CZSO	Diffe	rence
Goat categories	as at	as at	heads	%
	1 April 2022	1 April 2022	neaus	(100 = CZSO)
Live goats	48 152	24 607	23 545	+95,7
Goats - breeding females	30 052	16 661	11 887	+80,4
Other goats	18 100	7 946	10 154	+127,8

Commentary on tables 3. and 4.:

- The numbers of sheep and goats in the animal register are overestimated much more than the numbers of cattle.
- However, the reason is the same there are records of probably no-longer-alive animals and inactive farms (which is obvious from the too old date of birth of the animal, or in the age structure there are only animals with an old date of birth).
- The difference is even more striking also due to holdings/farms that do not reach the threshold for registration in the CZSO's Farm Register.
- The CZSO will negotiate with the MoA about the possibilities of eliminating such animals by setting appropriate conditions for pre-selection.
- The existence of a milk collection room in the holding/farm seems not to be a sufficient identifier of whether the sheep are milk sheep or non milk sheep. Respondents of the statistical survey *Livestock Survey* reported milk sheep also in farms without a milk collection room. This is especially true in case of farms where sheep's milk is obtained for own consumption.
- The CZSO will discuss with the MoA the search for a more suitable attribute or a combination of attributes that will have a sufficient rate of completing in the animal

- register and will be able to clearly determine the livestock yield focus of the holding/farm.
- Users of the IAR data use the age classification to determine the category of breeding female sheep/goats, i.e. that all females over 1 year of age are included in this defined group. Moreover, when a female sheep/goat has a registered delivery than is considered an ewe; without a registered delivery it is considered a breeding ewe-lamb/young goat. The definition of breeding ewe-lambs/young goats would then differ from the definition established for animal production statistics, where a breeding ewe-lamb/young goat is only considered to be an ewe-lamb put to the ram/a goat mated for the first time.
- The CZSO will ask the MoA to focus on increasing the completeness and quality of the data kept in the animal register (supervision of records of mandatory data).

B1. Requirements for External working group

The results of the analysis revealed problematic areas that need to be discussed and resolved with the Ministry of Agriculture in the upcoming period of the Grant Action:

- Registered no-longer-alive animals and inactive farms.

 The CZSO proposes to set up pre-selection conditions when extracting data from the animal register in such a way that the occurrence of records (animals, enterprises) that show signs of non-existence is eliminated to the maximum extent. These records are in the register because for formal reasons it was not possible to terminate them.
- Missing data.

 The CZSO will require that situations, where mandatory attributes (e.g. date of birth) and non-mandatory attributes (e.g. delivery of a female) are missing in the records, are minimized, especially for records that have met the pre-selection conditions.
- Insufficient differentiation of dairy animals from other animals.
 The CZSO will require the revision of the attributes needed to differentiate dairy animals.

B2. Output table preparation

The CZSO started preparations for defining the validation rules for checking the correctness of extracted data from administrative data sources (duplicities, logical data checks).

B3. Preparation of adjustments to the Statistical Metainformation System

The implementation of changes to the CZSO's Statistical Metainformation System is ongoing. Items in the Livestock code nomenclature were modified and new ones were added with respect to the hierarchical organization of items and the creation of aggregates according to user requirements.

B.4 Preparation of the solution for missing items

Testing of different models for providing missing (unavailable) items from the IAR data will continue in the coming period. If several models (for sheep) are available, the models used will be compared to select the one that would provide the most acceptable results.

- Crop production statistics

A.1 Analysis of ADS technical elements

Because of the analysis of administrative data sources (ADS) for the crop production statistics (LPIS and IS SAIF), the CZSO studied in detail the methodology, definitions and code lists used by the MoA and SAIF in connection with the administration of agricultural subsidies within the Single Application in LPIS. For the crop production statistics, there are two basic variables with the potential use of ADS, (a) the *Main area* and (b) the area of crops from arable land (*Sown area*).

- (a) Main area (Table 4 Agricultural land use in Regulation No 543/2009)
- Items UAA, arable land, permanent grassland and permanent crops are monitored in LPIS and in the crop production statistics.
- Definitions, based in national legislation (Agricultural Act No. 252/1997 Coll., on agriculture) and EU legislation, are applied in both evidences. There are no differences between the definitions.
- There are, however, differences in the published data between these two evidences for Orchards and Vineyards.
- Data on orchards are stored in the Orchard register administered by the Central Institute for Supervising and Testing in Agriculture (ÚKZÚZ), a state research institute subordinate to the Ministry of Agriculture. The Orchard register is a special register maintained under Agricultural Act and is linked to LPIS.
- In 2023, the MoA and the ÚKZÚZ plan a thorough modernisation of the Orchard register and the CZSO will closely cooperate in this activity with the aim to broaden the use of data for statistical purposes (especially the variables main area, production area and production by individual species).
- Data on vineyards are kept in the Vineyard register administered by the ÚKZÚZ pursuant to Act 321/2004 Coll., on Viticulture and Winemaking. The Vineyard register is in accordance with EU legislation. It also includes vineyards, which are not registered in LPIS because their owners do not apply for subsidies.
- The CZSO currently uses data on vineyard areas from the Vineyard register for the crop production statistics (main area, production area).
- (b) Area of crops from arable land (Table 1 *Crops from arable land* in Regulation No 543/2009)
- Data on sown areas of crops are collected by the CZSO in a sample statistical survey (according to Regulation 543/2009 on crop statistics).
- In the analysis of the availability of data from ADS, the CZSO already took into account the future requirements of SAIO regulation.
- Areas of crops are also required in the Single Application for subsidies, where farmers, farming on more than 10 ha of arable land, have to declare the area of grown crops as part of the greening (Payment for farmers following agricultural practices favourable to the climate and the environment).
- It is also mandatory to mark the crops in the respective Farmer's block (DPB) map in the LPIS.

Administrative data sources therefore contain data on the area (main area and area of crops from arable land) which can be used for the crop production statistics.

A.2 Analysis of user needs

- CZSO closely cooperates with the Crop Research Institute (VÚRV) in the field of nutrient balances (a working group *Nutrient balances*). Crop and animal statistics data are basis for the calculations of nutrient balances (N and P).
- The CZSO is also a member of the working group *Emissions* whose other members are the Crop Research Institute, the Research Institute of Agricultural Engineering (VÚZT), the Ministry of Agriculture and representatives of professional public and expert unions. This working group discusses the issue of greenhouse gas emissions.
- The CZSO informed both working groups and the professional public about the changes in the system of compiling agricultural statistics, where to a certain extent statistical surveys would be replaced by administrative data sources.
- Expanding use of ADS, the unification of methodologies and the prevention of double data reporting are all welcomed by the professional public.

A.3 Item/variable analysis

Analysis of legislative requirements

The detailed item-by-item analysis of administrative data sources was preceded with the analysis of the legislative requirements for crop statistics - according to Regulation 543/2009 on crop statistics and the SAIO regulation.

As a first step, the CZSO proposed an adjustment of the structure of crops surveyed in the crop production statistics for the year 2023 to be fully in accordance with the Crop Regulation and with the changes in the SAIO regulation, while reflexing also current changes in agricultural practice.

Based on this analysis the CZSO identified 7 items to be eliminated from the crop statistics from 2023 onward, because these items are either not required by EU legislation, or have only little economic significance, or data are provided by very low number of respondents (long term situation), or are not available from ADS.

List of items to be eliminated:

- Winter cereal mixtures for grain (small scale of growing, a low number of respondents, not included in the list of crops eligible for subsidies, not available in ADS, starting 2023 they will be marked as non-significant in the CROP statistics).
- Medicinal plants, cumin and culinary plants (currently available separate data for the three items are not possible in ADS; from 2023 they will be replaced with a summary item aromatic, medicinal and culinary crops in accordance with EU legislation).
- Gooseberry (small scale of growing, not required by EU legislation).
- Kitchen gardens (not available in ADS, not monitored separately in LPIS, not required in the SAIO regulation).
- Nurseries (until 2022, nurseries are included in arable land in the Czech crop statistics. From 2023 onwards, nurseries will be monitored separately in accordance with EU legislation and LPIS.

On the other hand, based on the verified extent of their growing and in accordance with the requirements of the current regulation and the SAIO regulation (especially organic farming production statistics), a list of 21 new items was identified to be included in the structure of crops grown on arable land, vegetables and fruits (permanent crops).

Crops from arable land (Annual crop statistics (ACS) Table 1):

- durum wheat
- buckwheat (new item according to SAIO)
- vetches (new item according to SAIO)
- winter rape seeds (replacing the summary item *Rapeseeds*)
- spring rape seeds (replacing the summary item *Rapeseeds*)
- aromatic, medicinal and culinary crops (a summary item replacing three separated items)

Vegetables (ACS Table 2):

Current summary item *Other vegetables* (monitored until 2022) will be replaced with the following list of vegetables to monitor the full list according to EU legislation:

- gourds and pumpkins
- watermelon
- asparagus
- other root, tuber and bulb vegetables
- other vegetables cultivated for fruit
- other fresh pulses
- other Brassicas
- other leafy and stalked vegetables

Permanent crops (ACS Table 3):

Similarly to vegetables, the list of monitored fruit species was supplemented with the following items:

- black currant (replacing the summary item *Currants*)
- red and white currant (replacing the summary item *Currants*)
- blueberries
- other pome fruits
- other stone fruits
- other berries
- other nuts

Analysis of ADS (LPIS, Single Application system)

The CZSO performed a detailed item-by-item analysis of the code nomenclatures used for crops in the LPIS in the administration of agricultural subsidies (the MoA is the administrator of these code nomenclatures).

Two crop code nomenclatures are used:

- Basic code nomenclature individual types of agricultural crops are listed (491 items).
- Aggregation code nomenclature (so called Central subsidy crop code nomenclature) includes 323 items that meet the needs of individual subsidy titles.

Crop production statistics of the CZSO are based on the division of agricultural crops according to the final product, which means that for cereals, legumes and oilseeds, the crops are divided into grain and green. However, in the subsidy administration system of MoA/SAIF,

only the species of crop is distinguished, not the use. This methodological difference means that without a fundamental modification by the Ministry of Agriculture (in LPIS) this source cannot be used for crop production statistics. The CZSO presented the results of this analysis to MoA and SAIF. During the follow-up discussions, it emerged that the direct modification of the Central subsidy crop code nomenclatures used in LPIS is not possible.

Instead, the MoA and SAIF proposed to add a new attribute *Purpose of growing the crop* in the crop declaration part in the Single Application administration in LPIS.

With the proposed adjustment (added new attribute) and other minor modifications to crop code nomenclature in the LPIS the CZSO assumes that this ADS can be used for crop production statistics.

A.4 The LPIS/Single Application system coverage

Farmers who farm on the area of at least 1 hectare of utilized agricultural area (UAA) can apply for subsidies using a Single Application. The threshold value for declaring crops in the Single Application is 10 hectares of arable land (or less, according to individual subsidy titles).

In 2022, 30,330 farmers, who applied for a subsidy in the Single Application, declared in total 3,534 thousand hectares of UAA. For comparison, in the Farm Register managed by the CZSO, in total 3,548 thousand hectares of UAA were registered. Data from ADS (Sigle Application) therefore cover 99.6% of UAA registered in the Farm Register.

Crop declarations in the Single Application - greening were submitted by 14,949 farmers who are farming at 2,469 thousand hectares of arable land. For comparison, 2,492 thousand hectares of arable land are registered in the Farm Register. ADS therefore covers 99.1% of arable land registered in the Farm Register.

In terms of coverage of UAA, this ADS meets the requirements set by EU legislation.

A.5 Updating of ADS (LPIS, Single Application system)

The update of records in LPIS is carried out by farmers continuously throughout the year, as stipulated in the LPIS methodology and in the relevant legislation. The farmer is obliged to report any change in the use of agricultural land in LPIS within 15 days.

The deadline for submitting the Single Application for subsidies is by 15 May, however, there is an additional deadline for "late" submission by 10 June (i.e. additional 25 days).

Based on the deadline for providing data on sown areas according EU legislation (30 June of year N), two dates were set for the extraction of data from ADS:

- 16 May of year N for **preliminary** data
- 15 June of year N for **final** data (no revisions are planned)

The term for final data corresponds to the current crop statistics, where sown areas are monitored by a sample survey as of 30 May and the resulting data are final, with no later revisions. The only exception is a later specification of the harvested area of maize, where most frequently occur changes in the final purpose of growing during the year – whether for grain (cereals) or as green maize (fodder). Similarly, the harvested area is also monitored for vegetables.

A.6 Data analysis

During 2022, a pilot project is underway to verify the functionality of the implemented changes in LPIS (as described in point B1).

Based on a bilateral agreement, the Ministry of Agriculture provided the Czech Statistical Office with three datasets to be subject of analysis:

- 1. The first dataset contains data from Single Applications for subsidies as of 15 May 2022. Farmers included in the dataset are those who decided to voluntarily participate in the pilot project. When completing a crop declaration, they also filled in the attribute *Purpose of growing the crop*.
 - During the inspection, however, it was found that the datasets contained duplicate records caused by the application submission methodology: farmers are allowed to fill in several applications and only later decide which option is the right one and forward it for further processing in the Single Application system. The MoA will work on adjusting the methodology and removing duplicates.
- 2. The second dataset contains data from Single Applications for subsidies as of 15 June 2022. The database contains data from 30,330 farmers who submitted the Single Application for subsidies in 2022 and declared the UAA. Data for the main land use category are available (arable land, permanent crops, permanent grassland).

Out of the total number of 30,330 units, 2,481 farmers participated in the pilot project (i.e. filled in the attribute *Purpose of growing a crop*). For these units, the sown area of grown crops is available.

The CZSO focused on comparing data for 1,733 units that were chosen for the CZSO sample statistical survey *Areas under Crops Survey* (a sample set of 10,000 units) and, at the same time, participated in the voluntary pilot project in LPIS.

Results of the comparison are presented in the following table.

Table 5. Sown area of the main groups of crops in 2022 (hectares)

Main groups of crops	Pilot project in Single Application –	CZSO sample survey	Difference CZSO/Pilot project		
	crop declarations	Areas under Crops 2022	ha	%	
Cereals for grain	417 250	433 615	16 365	3,9	
Dry pulses and protein crops for grain	15 207	14 332	-874	-5,8	
Root crops	23 501	23 556	56	0,2	
Industrial crops	141 055	141 124	70	0,0	
Plants harvested green	153 284	139 261	-14 023	-9,1	
Vegetables	1 659	1 493	-166	-10,0	
Strawberries	53	53	0	0,3	
Flowers and ornamental plants	1	5	4	232,1	
Seeds and seedlings	4 914	4 957	43	0,9	
Total	756 923	758 396	1 473	0,2	

Data at the level of individual units and items will be the subject of a deeper analysis the result of which develops the basis for modifying the mapping (see point B.1).

The following table shows the agricultural land use (main area) declared by farmers who submitted a Single Application for subsidies in 2022 (i.e. 30,330 units) and the data published by the CZSO in the crop statistics based on *Areas under Crops Survey*. The main differences are recorded in the main area of orchards and vineyards. The Vineyard register that is the data source for vineyard areas for crop statistics also includes vineyards not registered in LPIS. The CZSO will deal with the topic of orchards in 2023

Table 6. Agricultural land use declared in the Single Application and published by the CZSO in 2022 (hectares)

Main area	Single Application	CZSO	Difference CZSO/Single Application	
			ha	%
Utilized agricultural area	3 533 973	3 530 423	-3 550	-0,1
Arable land	2 493 136	2 484 911	-8 225	-0,3
Vineyards	15 060	17 448	2 388	15,9
Orchards	13 715	15 419	1 704	12,4
Permanent grassland	1 003 112	1 002 710	-402	0,0
Nurseries	1 997	1 910	-87	-4,3
Kitchen gardens (till 2022)	-	1 289	-	-
Other permanent crops	6 954	6 738	-217	-3,1

3. The third dataset contains data from the Single Application as of 15 June 2022 for farmers who filled out the crop declaration but not the attribute *Purpose of growing the crop*. Farmers with more than 10 ha of arable land must fill in the crop declaration in the application for subsidies (greening).

The database contained data for 14,949 units that declared 2,468,949 ha of arable land. However, the data do not correspond to the structure of crop production statistics as required by EU legislation (where the division into growing for green/grain/seed for cereals, legumes, oilseeds and fodder crops is required). The crop declarations without the completed attribute *Purpose of growing the crop* cannot be used for the purposes of crop production statistics. However, the CZSO will use these data to refine the estimates for the unsurveyed population during the further processing of crop production statistics.

In the following table, the Single Application crop declaration column shows the data from the Single Application for the greening without completed attribute *Purpose of growing the crop* for 14,949 units (without division into green/grain/seed) and the CZSO published data for crop production statistics in the other column.

The table shows, for example, a discrepancy in items Cereals for grain and Arable fodder crops. Both items include maize, but in the Single Application column, the total area of maize is included in plants harvested green from arable land (as if it were entirely grown for green, since no grain/green distinction is made). Similar discrepancies can be seen in the item Pulses and protein crops (grown for grain or green).

The analysis also confirmed that the Single Application is not able to provide data for the items Flowers and ornamental plants or Seed and seedlings.

Table 7. Sown area of the main groups of agricultural crops in 2022 (hectares)

Main groups of crops	Single Application – crop declaration (without purpose of growing)	CZSO published data	Difference CZSO/Single Application	
			ha	%
Cereals for grain	1 306 372	1 385 734	79 362	6,1
Dry pulses and protein crops for grain	52 327	45 634	-6 693	-12,8
Root crops	79 019	80 417	1 398	1,8
Industrial crops	447 800	444 977	-2 823	-0,6
Plants harvested green	545 604	467 363	-78 240	-14,3
Vegetables	12 365	11 678	-687	-5,6
Strawberries	406	444	38	9,3
Flowers and ornamental plants	23	219	195	843,2
Seeds and seedlings	4 787	15 270	10 483	219,0
Total	2 448 704	2 451 735	3 031	0,1

The CZSO will conduct a detailed analysis of the data during September 2022.

A.7 Study visit

Study visits to selected member states were not implemented during the first 12 months of the project. Partly due to travel restrictions (anti-covid measures), partly due to the fact that there was no need to consult any methodical procedures yet.

B.1 Requirements for external working group + **B.2** Output table preparation

The MoA, SAIF and the CZSO have agreed that the MoA will forward data from the ADS (LPIS, Single Application) to the CZSO for analysis purposes in 2022. Regular transfer of data (at the micro-data level) will be ensured by a contract between the owner of ADS and the Czech Statistical Office.

For the data transfer, the CZSO in cooperation with the MoA and SAIF defined the structure of the data. The data structure includes:

- a unique identification of the unit (a uniform identification code used in the Common Agricultural Register (SZR), an identification number (IČO), the name of the unit)
- a unique identification of the part of the Farmer's block (ID of the DPB)
- attributes linked to the DPB (the code for the type of agricultural crop, organic farming regime, the acreage of the DPB)
- an identification of the grown crop (according to code nomenclatures of the MoA and the CZSO)
- an attribute *Purpose of growing the crop* and the declared acreage of the crop

Structure of the data:

Table 8.

Group	Description of attributes		
	LPIS ID of the user		
	SAIF Uniform identifier of the user		
User identification	SZR ID of the user		
	Identification number (IČO) of the user		
	Business name of the user		
	ID of DPB (part of the Farmer's block)		
DPB identification	Short code of DPB		
Dr B identification	Square code of DPB		
	Code of the prevailing cadastral territory		
	Agricultural crop code		
Attributes linked to	Organic farming regime		
DPB	Prevailing area with natural or other special restrictions		
	DPB acreage [ha]		
Identification of the grown crop and its acreage	ID of the grown crop (in the Central subsidy code nomenclature)		
	Name of the grown crop (in the Central subsidy code nomenclature)		
	Purpose of growing the crop		
	ID in the CZSO survey questionnaire - crop code nomenclature (208, 209)		
	Item name in the CZSO survey questionnaire		
	Row number in the CZSO survey questionnaire		
	ID code of the drawing of the parcel with the declared crop in LPIS		
	Declared acreage in the Crop diversification measure [ha]		

In the first phase, the MoA's Central subsidy crop code nomenclature was modified. Items necessary for the crop production statistics were added. The option to tick the attribute *Purpose of growing the crop* was added to the declaration of crops to state the purpose of growing (for grain/green/seed/other product).

In the second phase, the CZSO in cooperation with the Crop Research Institute created a converter (mapping) between the items (crops) listed in the MoA's Central subsidy code nomenclature (323 items) and the CZSO's crop nomenclature (82 items). The attribute *Purpose of growing the crop* was defined for each crop or group of crops including setting the default prevailing purpose.

In the third phase, based on the mapping, a special output (composition) for farmers called *Report on sown areas for the CZSO* was programmed in LPIS. When a farmer declares crops for subsidy purposes in LPIS, he confirms the default attribute *Purpose of growing the crop* (or chooses the correct one) and the data on declared crops are automatically transferred to items of *Report on sown areas for the CZSO*. In the following step, the farmer confirms the correctness of filled sowing areas for individual crops. The MoA then forwards the database to the CZSO in the agreed structure and terms (see above in A.5).

In 2022, a pilot verification of the functionality of the above described changes set in LPIS was carried out. Participation in the pilot verification was voluntary; 2,481 farmers took part in the pilot study.

B.3 Preparation of adjustments to the Statistical Metainformation System

The implementation of changes in the Statistical Metainformation System (SMS) of the CZSO is ongoing. CZSO has started work on the modification of the SMS for the crop production statistics, which includes:

- the refining of definitions in code nomenclatures,
- translation of definitions into English,
- modification of statistical questionnaires,
- modification of the methodology for data collection and processing,
- definition and modification of data checks and time schedules,
- preparation of output tables.

B.4 Preparation of the solution for missing items

Based on the analysis results, several items not available in ADS were identified. Alternative data sources need to be found. These items are:

- Flowers and ornamental plants (proposed solution: A supplementary survey to the sample survey *Areas under Crops Survey*, the sample to be defined from the Farm Register).
- Christmas trees (proposed solution: A supplementary survey to the sample survey *Areas under Crops Survey*, the sample to be defined from the Farm Register).
- Vegetable and strawberries under glass or high accessible cover (proposed solution: A supplementary survey to the sample *Harvest Survey*, the sample to be defined from the Farm Register).
- Seed and seedlings (proposed solution: Use of other ADS, the certification procedure of seeds and seedlings and registration of propagation stands in the Central Institute for Supervising and Testing in Agriculture).

B5 Methodology for updating the Farm Register

CZSO performed an analysis of administrative data sources (LPIS, Single Application system and the animal register), from the point of view of their structure (items and attributes). Due to the appropriately defined structure of the data sentence, the structure of the transmitted data was evaluated as satisfactory for the needs of updating the Farm Register and no adjustments to the methodology had to be introduced.

Analysis in terms of item identifiers is still ongoing. The preliminary analysis shows that it will be necessary to develop new methodologies for updating units without identification number of the user (IČO).

B6 Update of the methodology for linked statistical surveys

In order to facilitate the process of transition from statistical surveys to administrative data sources, the entire process of data extraction from the ADS and its use for statistical purposes was maximally adapted to the currently used methodology of generating samples from the Farm Register for statistical surveys and to the current schedule of all operations in frame of the preparation of statistical surveys. Apart from the *Livestock survey*, no other existing sample statistical survey will be cancelled, but only reduced.

The current methodology for updating the Farm Register for CZSO's statistical surveys was evaluated as fully sufficient. In the next phase, CZSO will analyze suitable solutions allowing increasing the number of attributes extracted from ADS and thereby also cover marginal areas of agriculture (e.g. individual types of vegetables).

Further work envisaged

- In the period evaluated in this interim project, the pace of work followed the planned timetable, and therefore it is not necessary to adjust the timetable.
- Cooperation with other institutions (External working group) was excellent and further negotiations will continue on the points listed below.
- In the next phase of the Grant Action the work will focus on:
 - The CZSO will continue the negotiation with the MoA regarding the minimizing of factors limiting the use of IAR data for statistical purposes (inactive farms, overaged animals, missing mandatory and non-mandatory data), increasing the quality of data and the completeness of the animal register.
 - For sheep and goats, suitable attributes for the determination or modelling the numbers of breeding females and the distinction between milk/non milk breeding females need to be found. The negotiation of the CZSO with the MoA is ongoing.
 - Testing of different models for providing missing (unavailable) items for sheep.
 - Detailed analysis of data from the pilot project (crop statistics).
 - Editing of the mapping of crops, in cooperation with the Crop Research Institute.
 - Modifications of requirements for the Ministry of Agriculture for crop statistics, statistics of permanent crops and the Orchard register.
 - Alignment of LPIS methodology and the methodology of the crop production statistics.
 - Formulation of the methodology for the sample survey *Areas under Crops* (sampling methodology for this survey and for linked surveys).
 - Data processing (grossing-up).
 - Output tables (output programming in the Statistical Metadata System).
 - Dissemination of results in the online database on the CZSO website (definition of tables).
 - Work on the Farm Register (see B.5 and B.6 above).

In the fourth quarter of 2022, the CZSO, the MoA and SAIF will discuss a framework agreement on the transfer of administrative data for statistical purposes. Subsequently, the contract will be supplemented with appendices defining detailed requirements for the data to be transferred following the results of the works listed above.

In August 2023, a second interim Progress Report will be submitted to Eurostat.