MULTI COUNTRY STATISTICAL COOPERATION

Eurostat Pilot Projects on Statistics

Financed by EU, Phare programme

MULTI COUNTRY STATISTICAL COOPERATION

PILOT PROJECTS ON STATISTICS:

COMPONENT: AGRICULTURAL SECTOR MODELLING

Use of FADN data as a basis for the generation of data on intermediate consumption in Latvia

Final report

Liene Dambina, Valda Bratka, Andris Miglavs,

Oskars Rasnacs, Ruta Filmanovica, Signe Zvirbule

Latvian State Institute of Agrarian Economics

2001

Content

Tables

INTRODUCTION

This study is prepared under the Contract "Provision of Services from the Phare Countries" in the framework of "Multi Country Statistical Cooperation - Pilot Projects on Statistics", Component: Agricultural Sector Modelling.

The objective of the study is to develop and test an approach for using data from the Farm Accountancy Data Network (FADN) and other data sources for generation of input use coefficients per Agricultural Activity[1].

The work is based on 2 principle data sources:

- Farms' data from the Latvian FADN database, developed at the LSIAE,
- The data on input coefficients available from existing ASM-data bases in Latvia (at the LSIAE).

All the calculations done are based on data from 1999 - the last year for which data from all sources are available.

The tasks of the study were formulated as:

- to outline an approach for continuous generation of input data per activity in Latvia;
- to develop recommendations for approaches to be applied in other CC's.

According to the ToR, the objective of this sub-project is not to be a one-time (ad-hoc) exercise, rather to elaborate a methodology which can be applied regularly and which would provide a regular data flow to Eurostat concerning the input coefficients.

Eurostat has proposed a minimum level of detail of the inputs, which is in line with the level of detail in the EAA and AgrIS. This include **input data** about:

- seeds,
- energy (electricity, gas, other),
- fertilizers (manure, purchased),
- plant protection (pesticides),
- veterinary expenses,
- animal feed (farm produced, supplied by other agricultural holdings, purchased outside agriculture),
- maintenance of materials and buildings
- agricultural services;
- other goods and services.

If more detail were available on for example fertilizers (nitrogen, phosphate, potash) this would be favourable.

The study was carried out in several stages:

- In the first step the existing set of input coefficients was examined. Focus was put on the approaches currently applied for generating these data, including the evaluation of their consistency to EAA and other statistics.
- Based on the above, the methodology was elaborated and tested how Latvian FADN-data can be used to upgrade the set of input coefficients, and the issues were outlined.
- Finally, a set of input use coefficients per agricultural activity was generated and an analysis of data reliability was carried out, calculating and evaluating the difference

[1] Later in text - input coefficients

between the desirable level of costs per sector, calculated as generated input coefficients multiplied with the relevant activity level on the one hand, and the value of the correspondent cost item according to.

This report represents the findings of the study, and consists on three major chapters, according to the stages of the work done. Also the concluding part, where the summary and main conclusions are presented, is added.

From all the activities and cost items only those covered by the analysis of this study are discussed in the report.

1. CURRENT SITUATION IN LATVIA

The Latvian State Institute of Agrarian Economics (LSIAE) carries out the development of Economic Accounts for Agriculture (EAA) and its gradual implementation in Latvia since 1994. From the European Union (EU) side, experts from the ASA Institute of Sector Analyses and Policy Advice in Bonn, Germany were the main advisers during all this time.

Currently, LSIAE is working on establishing an integrated database, which could find broader use for policy analysis purposes. In the course of that work the contacts with the Central Statistical Bureau (CSB), the Latvian Agricultural Advisory and Training Centre (LAAC), the Latvian Agricultural Trading Agency (LATA) and other institutions collecting data have been established and suitability assessment of the available information for EAA needs has been performed.

The main data source for EAA is Central Statistical Bureau (CSB). The agricultural unit of the CSB provides most of the information. According to the CSB structure of the data collection, there are 2 groups of data:

- covering the entire agricultural sector (e.g. data collected from processing enterprises about the sales of products).
- covering activities of the main groups of Latvian agricultural producers: agricultural companies and state farms, on one hand, and individual - family farms.

Information from individual (family) farms is obtained from sample survey taking place twice a year – in June and in November. In 1999 it included 13 000 farms or approximately 10 % of farms with at least 1 ha of utilized agricultural area or producing specific products for sales on less area. The sample of the survey has been designed on the basis of agricultural holdings Register maintained by CSB. Although we should admit, data from these surveys not always are based on accountancy, and especially the economical (financial) results must be treated with some reservation.

Statistical reports of agricultural companies are collected directly and, mostly are based on accountancy data. It gives some more inside information about their economy. However, as the weakness of this data source, the limited representativeness for the whole sector should be outlined, because the share of this type of farms (agricultural companies) in total agricultural output has considerably decreased during the restructuring process, taken place in last decade. In 1999 it was just 23 %.

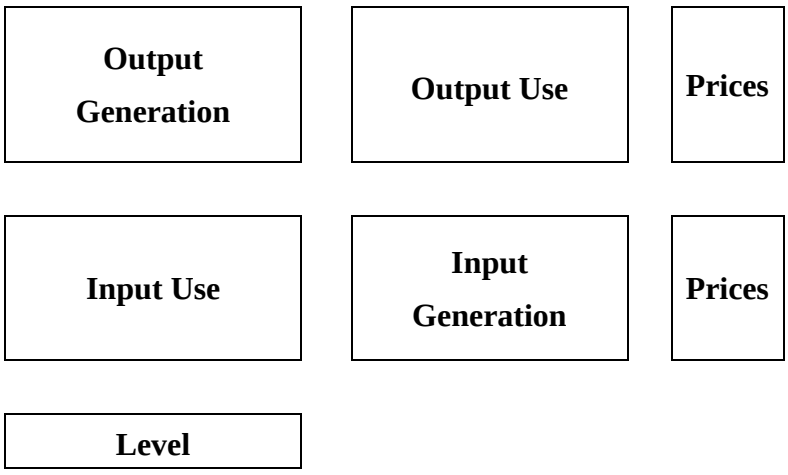
Some information is taken also from other sources: for example, prices for the valuation of self – produced stock are obtained from the LAAC.

1.1. Approach and structure

The activity - based accounting approach (ABTA) is used for EAA calculation in Latvia and an EXCEL tool applied. The Accounting framework is composed of 5 main blocks: Output Generation, Output Use, Input Use, Input Generation and Level (crop area and livestock number per activity). Prices are forming a separate block in this framework (Figure 1).

Thus in Latvian ABTA the columns are representing the **production activities** (output and input generation and use). Rows are representing the **products** and **input items** according to the different activities.

Figure 1. ABTA accounting framework



The structure of this approach enables consistency and plausibility checks since it covers the total agricultural production by supply and demand, and by outputs and utilised inputs. Consistency checks compare the amount of supply (production statistic) and the aggregate amount of market and non-market use (supply balances) for each product. Similarly the aggregate input use by production activities has to be equal to the domestic supply of agricultural inputs.

1.2. Current input coefficients

The detailed description of the input use per each crop and livestock production activity is not available in statistics. CSB can provide input use data only for the sector in total by the main input items, because only the data about the input use in the whole sector are available from statistical reports.

In 1995 as the main data source for input use coefficients the special studies, called "Analytical reviews by sectors", were used. These reviews were made by the groups of researchers, established to work out the concept of further development of Latvian agricultural sector. In the framework of those studies also the input use levels per different activities were investigated. In 1996 these coefficients were updated using gross margin calculations done by LAAC, based on the data from LAAC demonstration and bookkeeping farms.

For 1999 the input use coefficients were updated again, based on some more recent farm surveys. More detailed description of used data sources and expert estimations is given in the next chapter.

1.2.1 Data sources and calculation

Seeds and planting stock

Data on grain, potatoes and vegetables are obtained from the CSB farm survey data on consumption of self-produced inputs in individual farms, while data about the self-produced and purchased input use in agricultural companies and in state farms was taken from their annual statistical reports. Concerning the use of seed in 1999 calculations the data from Ministry of Agriculture, calculated according to average standards, have been used. The use of self-produced seed has been valuated in internal consumption prices (depending on type of product, they were assumed as 90-100% of farm gate price). Purchased seed is accounted in farm gate prices.

For the rest of products seed consumption standards from LAAC gross margin calculations have been used and prices have been assessed by the expert method, based on different available data sources.

Petrol and lubricants

There are data from LAAC gross margin calculations used in approximation of expenditures on petrol and lubricants in crop sector. The costs of agricultural services also are taken from these data. There were special coefficients calculated by LSIAE describing the share of petrol and lubricants in the total costs of agricultural services.

Cost calculations from the development programmes for dairy, beef and pork sectors have been used as input data source for the livestock production.

For the rest of activities in livestock sector the costs have been assessed using the expert method.

Electricity, fuel for heating

Data from LAAC book keeping farms, which are specialised in a particular production activity (grain, potatoes, sugar beet, vegetables, dairy, beef, pork), are used in the calculation of expenditures on electricity in crop sector. For other activities, the calculations, which were done for the sector development programmes, are used, or, alternatively, data from similar activities, applying the expert method, were used.

Fertiliser and soil improvers

Use of fertilisers by crop production activity is obtained from CSB survey, which covers state farms and agricultural companies and reports consumption of fertiliser (in pure substance) and manure by type of activity in particular cropping year. For individual farms the use of fertiliser is report only as total figure. To get specification, the proportions between the different types of substance calculated from the companies' survey are applied. The price of fertiliser is obtained by dividing the total costs for fertiliser and soil improver (less purchased manure, the share of which is assessed by expert method) by the amount of pure substance in the used fertiliser. The price proportions between elements N, P and K are the same as used by LAAC 1:4:2.5.

Plant protection products

Data on consumption of plant protection products by activity are obtained from pesticide consumption normative, reported in LAAC gross margin calculations, where they are given separately - for farms with intensive and extensive production. Average indicators for the whole activity are calculated as average weighted, where the proportion of high and low yield areas (reported in CSB Structural Survey) are taken as weights (based on expert assumption about the yield threshold to be counted as intensive production)

Veterinary expenses

Data about veterinary expenses on one animal or fowl are taken from LAAC gross margin calculations. Data about the numbers of animals and fowls is taken from data of CSB about farms in Latvia.

Feedingstuffs

Data of consumption and prices of feeding stuffs are taken from LAAC gross margin calculations.

Maintenance of materials and buildings

Data about equipment and buildings maintenance costs in crop sector are taken from LAAC gross margin calculations like expenditures on services. There is a special coefficient, calculated by LSIAE, which describes the share of maintenance costs in the expenditures on agricultural services.

For livestock production cost calculations (dairy, beef and pork sector) "Analytical reviews by sectors" have been used as data source; for the rest of activities in the sector the expert method has been applied.

Labour costs

Labour costs are based on labour consumption standards developed for activity development programmes; where not available, data for similar production activities are used. An average labour cost 0.65 LVL/h was assumed for 1999 (including social costs).

Rent

In most cases the level of land rent payments is taken from FADN database, in particular- from farms with specialisation in crop production activities. To evaluate rent payments on fodder areas, FADN data about rent payments in specialised dairy farms were used.

Interest paid

There are data from LAAC gross margin calculations used in approximation of interest paid in *crop production activities*. There were special coefficients calculated by LSIAE describing the share of interest paid in the total costs of agricultural services, for which the rate per ha is given form most of crop production activities. Coefficients are calculated based on estimated capital costs, share of credits in financing the purchases of capital goods and average interest rate available for producers.

In livestock production the values of interest paid per activity unit are taken from FADN data analysis - data from farms with specialisation in particular activity.

1.2.2 Summary of current input use coefficients

Table 1. Input use coefficients per head of livestock production activity

Input items	Dairy cows	Fattening cattle	Fattening pigs	Sheep and goats	Laying hens	Poultry for slaught.	Breeding pigs	Adjustment coefficient
Petrol, lubricants	2,44	1,49	1,07	0,46	0,02	0,02	1,83	1,281
Electricity, fuel	10,00	3,46	2,05	1,14	0,61	0,42	4,09	0,868
Veterinary expenses	9,92	1,12	0,79	0,96	0,42	0,13	20,70	0,932
Feedingstuffs	154,99	87,62	63,76	34,19	2,85	1,50	165,67	0,915
Maintenance	8,27	3,20	0,99	0,76	0,03	0,03	1,99	0,335
Compensation of employees	25,28	8,00	1,58	1,58	0,16	0,16	2,05	0,228
Interest paid	1,65	0,43	0,25	0,25	0,00	0,00	0,25	0,125

Table 2. Input use coefficients per ha of crop production activity

Input items	Winter wheat	Spring wheat	Rye	Barley	Oats	Other cereals	Pulses	Potatoes	Sugar-beet	Rape	Flax
Seeds and planting stock	15,99	15,99	12,78	12,94	12,84	12,15	12,51	156,37	65,00	11,14	4,89
Petrol, lubricants	16,46	14,90	14,90	15,38	16,37	16,37	17,54	31,56	36,81	19,47	25,99
Electricity, fuel	1,35	1,35	1,28	1,28	1,16	1,28	0,97	1,35	0,70	0,43	0,55
Fertilisers and soil improvers	41,25	23,98	35,73	20,45	18,90	36,43	18,39	83,38	66,33	36,18	83,56
Plant protection products	25,00	31,20	13,00	13,00	8,50	13,00	34,00	28,27	102,00	39,00	45,00
Maintenance	12,67	11,46	11,46	11,84	11,84	12,67	13,50	24,29	28,32	14,98	15,77
Compensation of employees	28,60	20,80	26,78	19,24	20,28	19,24	28,60	84,50	106,73	23,92	74,23
Rent	3,88	3,88	3,88	3,88	3,10	3,88	3,88	3,88	3,88	3,88	3,10
Interest paid	34,84	31,53	31,53	32,55	32,55	12,67	37,11	66,79	77,89	41,19	43,36

Table 2 (continuation) Input use coefficients per ha of crop production activity

Input items	Other industrial crops	Vegetables in open areas	Vegetables in covered areas	Fresh fruits	Root crops	Crops for silage	Maize for silage	Hay	Other fodder crops	Adjustment coefficient
Seeds and planting stock	11,00	396,72	986,40	89,03	65,00	0,02	24,00	0,08	3,99	0,762
Petrol, lubricants	25,99	9,08	500,42	15,59	21,42	10,15	10,15	8,76	2,19	1,281
Electricity, fuel	0,55	0,80	22663,45	0,40	0,70	1,10	0,55	0,55	0,55	0,868
Fertilisers and soil improvers	20,45	12,35	120,21	4,14	53,70	5,47	46,38	5,62	3,67	0,969
Plant protection products	15,00	70,40	337,20	168,00	8,50	0,00	6,50	0,00	0,00	0,329
Maintenance	15,77	22,91	953,15	14,90	28,32	7,81	6,40	6,74	1,69	0,335
Compensation of employees	74,23	245,38	4653,76	65,00	104,13	6,21	7,87	9,22	6,93	0,228
Rent	3,88	3,88	3,88	3,88	0,59	0,59	0,59	0,59	0,59	0,513
Interest paid	0,00	66,26	355,00	113,76	6,74	21,48	21,48	18,53	4,64	0,125

2. USE OF FADN DATA FOR COST DISTRIBUTION PER ACTIVITY

The assignment is to allocate farm costs to individual farm production activities, using data from Latvian Farm accountancy data network (FADN). Cost distribution - calculation of inputs coefficients has been done based on Latvian FADN database for 1999 containing 520 farms data.

2.1. FADN farm sample in 1999

Setting up of farm account statistics (SUDAT) based on the principles of EU Farm Accountancy Data Network has been started in Latvia in 1996. It is a joint project of several parties: Ministry of Agriculture (MoA), LAAC and LSIAE as the leading institution agreed on joint activities to build up the SUDAT. LAAC is responsible for data collection, data check and providing to LSIAE, LSIAE - for data processing and publishing and also for the methodological background and development of the system. MoA has to provide necessary governmental support and legislative background.

First publication was prepared for 1996 based on the 222 LAAC bookkeeping farms data; for 1997 number of farms has been increased to 398 and for 1998 to 825: as an additional source to LAAC bookkeeping farms, data obtained from farms applying for subsidies were used.

Sample of 500 farms for SUDAT has been designed in the spring of year 2000. Farm sample has been designed based on three criteria: type of farming, economical size and region. Due to lack of farm register containing all necessary information for SUDAT: sown areas by different type of crops and number of animals (only data about total agricultural land operated by farm are available in farm registered maintained by CSB) sample has been designed based on the sources of information available at current stage: data from agricultural producers - legal entities and data from 13 thousand farms obtained from farm structural survey.

Farms are classified by three criteria compatible with EU: type of farming, economical size and region. Type of farming and economical size has been based on standard gross margin concept at current stage calculated only for the whole Latvia. Regional breakdown is based on NUTS level 3. Farm business analysis has been performed by regions, by utilized agricultural area, type of farming and economical size of farms.

Only 222 farms out of selected 500 were able and agreed to submit the data for 1999. With a design to have certain number of farms at SUDAT for 1999 as a temporary solution has been accepted to include data from 298 LAAC bookkeeping farms (as an additional source of information).

Therefore the farm account results for 1999 could not be considered as completely representative for Latvian agriculture, but they are able to give some inside view on processes taking place in Latvian agriculture.

2.2. FADN data

2.2.1 Structure and content of Latvian FADN database

Latvian FADN database has several parts and contain the following data:

- **General data** about the holding,

- **Farm land**: own land, rented agricultural land and land rented out;
- **Labour input**:
  - data about farm manager and holder;
  - annual working units and worked hours in agriculture, forestry and other farm businesses for paid and unpaid labour, regular and casual;
- **Number and value of livestock** at the beginning and at the end of year, average number of livestock by categories:
  - cattle: calves for fattening under 5 months, other cattle less than one year old, male cattle from one to less than two years old, female cattle from one to less than two years old, male cattle two years old or more, breeding heifers, heifers for fattening, dairy cows, cull dairy cows, other cows,
  - pigs: piglets, breeding sows, pigs for fattening, other pigs,
  - sheep: ewes and other sheep,
  - goats, breeding females and other goats,
  - poultry: table chickens, laying hens, other poultry,
  - rabbits, breeding females,
  - horses,
  - other animals,
  - beehives;
- **Movement of livestock**: purchases, birth, sales and farm household consumption by groups pf livestock: cattle, pigs, sheep, goats, poultry, rabbits, horses, other animals, beehives;
- **Production data** by types of crops and livestock products (except data on meat and live animals counted above), processing of agricultural products:
  - use of the agricultural land by crops (except livestock products and processing),
  - opening valuation: stocks and values of products,
  - production for the accounting year: quantities,
  - sold quantities and sales values,
  - quantities and values of farm use: for seeds, for feeding, for processing,
  - quantities and values of farmhouse consumption;
  - closing valuation: stocks and values of products,
- Data about **other farm receipts:** rural tourism, forestry and other farm businesses;
- **Farm costs**:
  - labour and machinery costs: wages paid and social security, contract work, current upkeep of machinery and equipment, motor fuel and lubricants,
  - specific costs of crop production: purchased seeds and seedlings, fertilizers and soil improvements, crop protection products, other specific crop production costs,

- specific costs of livestock production: purchased concentrate feed for grazing livestock, other purchased feed for grazing livestock, purchased feed for pigs, purchased feed for poultry and other livestock, insemination and veterinary fees, other specific livestock production costs,
- specific forestry costs,
- other farm businesses specific costs,
- farming overheads: upkeep of land improvements and buildings, electricity, heating fuels, water supply, insurance, environmental tax and other dues, other farming overheads;
- land charges: rent paid, insurance for farm buildings, taxes on land and buildings,
- interest paid: on loans for land and buildings, on loans for machinery and equipment, on other loans,
- extraordinary items,
- income tax on farming activities;

- **Subsidies and state support**: subsidies for crop production, subsidies for livestock production, general subsidies for farm, compensation of excise tax;
- **Fixed farm assets**: agricultural land, forest land and standing timber, buildings, machinery and equipment, circulating capital, financial investments and acquisition costs for quotas and other rights (opening and closing valuation, investment, sales, depreciation);
- **Farm taxes**: opening, closing, calculated and paid value;
- **Farm liabilities:** opening, closing, increase and paid back value**.**

2.2.2 Preparation of FADN data for analysis

The total farm costs to be allocated by activity and used as input data in this exercise are:

- pesticides,
- veterinary services,
- seed,
- feed,
- electricity and heating,
- fuel and lubricants,
- maintenance of machinery and buildings,
- labour,
- rent,
- interest.

Also areas and numbers of animals under different activities are known. Proceeding from these input figures our assignment is to calculate unit production costs, namely, per 1 ha of sown area and/or per 1 animal.

In principle, some of the farm costs could be directly allocated to particular activity or group of activities.

For example, farm grown seed by type of crop could be taken directly from production data about the farm use of production for seed (quantity and value). However in our study

still the total seed costs are distributed by activity, because the costs for purchased seed (counted as a separate item in farm costs) are dependent on the use of farm grown seed. And additional investigation on relationship between the consumption of self- produced and purchased seed would be needed, which is let for later studies.

Farm accountancy data contains separate cost items also for purchased concentrated feed for grazing livestock, other purchased feed for grazing livestock, purchased feed for pigs, purchased feed for poultry and other livestock, which could be directly allocated to particular animal group. The same groups of animals can use however the farm produced feed as well. Therefore the share of purchased feed in total volume of used feeding stuffs depends on farmer's strategy, which may be different and behaviour of an "*average farmer"* is not known. Due to that in the current study the total volume of feed costs was calculated first and then it was distributed according to the approach described later.

As the result the initial data for further use in the study was prepared, which covers all the above-mentioned activities and cost items to be allocated by the activities.

2.3. Methods

2.3.1 Multiple linear regression analysis – the principal method used

To solve this problem, multiple linear regression analysis method has been selected. According to the theory, multiple regression is a statistical method for analysing the relation between several independent (factorial) variables and one dependent variable. Independence of all variables except one dependant variable to which linear relationship from others exist is a precondition for using multiple linear regression analysis.

All factorial variables - planted areas under different types of crops and number of animals, in this exercise are independent, the resultant variable - total amount of cost item in the farm, is dependent on the factorial variables. This is the reason for selecting multiple linear regression method. It is used also in some other countries to solve similar problem, for example, in Denmark.

For each farm and for each cost item the following equation is true:

$$y=a_0+a_1x_1+a_2x_2+\ldots+a_nx_n,$$ where

y is the resultant variable, depending on the factorial variables x_1, x_2, $\ldots x_n$ and a_0, a_1, $a_2,\ldots,a_n$ are coefficients of regression equation to be calculated.

The factorial variables x_1, x_2, $\ldots x_n$ in our case are the number of animals (in the case of specific livestock production costs) and / or planted areas (specific crop production costs or both for farming overheads and fixed costs), particular cost item is assumed as resultant (dependent) variable.

Applying multiple linear regression analysis method, we should find a set of coefficients a_i , where $i=1,2, \ldots, n$ which meet the equations for all the farms in the best possible way.

The obtained regression coefficients a_1, $a_2,\ldots,$ a_n specify costs per 1 ha or per 1 animal; a_0 is an intercept. These costs may not be negative by definition: in the opposite case the economic meaning would be reproduction (generating) of inputs in the production process instead of their consumption. Therefore it is necessary to get non-negative or, in the worst case, close to zero coefficients in terms of absolute value. Intercept may have any value. In case of negative intercept the amount of allocated cost item is larger than actually used. If the intercept is positive, it means that not all the costs are related to the particular

activities. Having in mind the task to find the relationship between the input use coefficients, in principle the both results could be acceptable.

Therefore in this research work both the cases have been analysed, and the best result assessed by the sum of standard errors $\hat{s}$ in square was chosen. By calculating the sum of the standard error squares of regression coefficients, we can identify which of the results are more correct (reliable). The smaller the sum, the more correct (reliable) the result.

In this exercise the multiple linear regression tool offered by Microsoft Excel was used. It gives also a possibility to calculate standard errors $\hat{s}$ of regression coefficients and determination coefficient R^2. The determination coefficient demonstrates how well the equation fits with the data – how many percent of the points ($x_{j1}; x_{j2}; \ldots x_{jn}; y_j; j=1, 2, \ldots k$, where k – number of sample farms) satisfy regression equation. Standard errors determine the possible bounds for the calculated coefficients.

Microsoft Excel allows to do regression analysis if the number of factors is less than or equals to 16, although Microsoft Excel linear regression analysis tool does not allow applying constraints to avoid negative figures and apply break-even points to take into account economy of scale. If the number of factors exceeds 16, other tool, for example, regression analysis with SPSS, SAS should be applied. It was the first attempt to use FADN data for input coefficients calculation, but to obtain more correct data some other tool should be used.

2.3.2 Types of costs and the approaches of regression analysis

From the point of view of applicability of cost items to an activity, three groups of costs can be singled out:

1) related only to crop or livestock production;
2) related to both, crop and livestock production;
3) related to agriculture and also to other farm businesses, like forestry, processing, etc.

Depending on attribution of a cost item to one of the groups, one of the two main analytical approaches was chosen:

- one-stage regression analysis, where the amount of cost item was directly distributed by particular agricultural activity (the 1st and 2nd group of costs);
- two-stage regression analysis, where at first regression coefficients for different farm businesses were identified, then the amount of costs attributable to agriculture was calculated and finally this amount was distributed by activity (the 3rd group of costs).

As sub-variants simple regression and *Activity-wise cost analysis* can be outlined.

Let us consider these analytical approaches in more detail.

A One-stage regression analysis

The costs of pesticides, seed and fertiliser are applicable only to crop production. The costs of the veterinary services and feed are applicable exclusively to the livestock production. Labour consumption applies both to crop production and livestock production.

Planted areas are the factorial variables for costs applicable exclusively to crop production. Regression coefficients estimate the respective costs per 1 ha.

The number of farm animals is the factorial variable for costs applicable exclusively to livestock production. The regression coefficients estimate the respective costs per

1 animal.
Number of the animals and planted areas are the factorial variables for costs applicable to both agricultural production branches. In this case regression coefficients estimate the respective costs per 1 ha of crops and per 1 animal.
One - stage linear regression analysis is used for all the above cases. It means, that regression analysis tool is applied on the preliminary selected data set. The obtained regression coefficients are accepted as the input use coefficients per unit of activity.

B Two-stage regression analysis.

There are several costs applicable not only to agriculture (crop and livestock production), but also to forestry, processing, rural tourism, other industries. Electricity and heating, fuel and lubricants, maintenance of machinery and buildings, rental costs and interest belong to this group.
In this case a two-stage regression analysis is used. At the first stage the amount of costs applicable to the agricultural industry was identified. To do this, as the 1st step the regression analysis was applied, where the output of different industries in monetary terms was used as the factorial variable.

The second step was to calculate the part of costs attributable to agriculture by subtracting non-agricultural part, which could be calculated according to the following equation (residual principle):

$$y_{agr} = y - a_0 - a_2 * x_2 - a_3 * x_3 - a_4 * x_4 , \qquad \text{where}$$

x_2, x_3, x_4 are value of output of forestry, processing and other farm businesses;
b_2, b_3, b_4 - regression coefficients characterising the input use level per output unit in the industry (in monetary terms).

At the second stage we apply the approach described above -One-stage linear regression analysis. It means, that planted areas and the number of animals now are the factorial variables again.

C "Activity-wise cost analysis".

If application of *one-stage* or *two-stage* regression analysis lead to unrealistic results, "Activity-wise cost analysis" was tried. Farms, having a particular activity *i*, were selected and regression analysis for these farms was done. The regression coefficient b_i of the activity *i* was regarded as the result and taken for the final set of the coefficients. This analysis was repeated for all the activities regarding particular cost item.

This approach was applied for allocation of labour costs and fertiliser costs.

2.3.3 Application of One-stage multiple linear regression analysis

Pesticides costs

Application of *One-stage regression analysis* was carried out in several steps:

1) At first, we have to select farms for which the analysed cost item is not a zero.
2) We have to identify the activities, for which area for all the farms is zero or on which pesticide (or other cost item) costs are not applicable. The columns with a zero value for planted areas (and/or number of animals in other cases) should be discarded. In our case fungus, meadows and pastures were discarded.
3) To avoid the number of activities more 16 and if the areas under the particular activities are small, we can try to aggregate some activities with similar level of

input use. It was possible due to similar production technologies. In the case of pesticides we may group grain with leguminous plants, vegetables with strawberries.

4) Then regression analysis is done for all the selected farms and activities under the assumption that the intercept is 0. Thus we obtained the preliminary results for pesticides (Table 3).

Table 3. The results of the first regression analysis for the pesticides

Crops	Regression coefficients
Wheat	13,66
Rye	6,25
Barley	9,34
Oats	10,71
Other grain and leguminous plants	14,55
Potatoes	74,00
Sugar beet	126,72
Oil and fibre crops	14,56
Other field crops	-22,54
Field vegetables	21,40
Vegetables in the specialised areas and strawberries	62,36
Vegetables in covered areas	-28524,61
Flowers and decorative plants	-3066,82
Flowers and decorative plants in covered areas	124945,09
Fodder crops	0,85
Perennials	-10,15

5) We have to identify the existence of factors (crops and animals), which have regression coefficients with either negative or very high positive value. Regression coefficients may not be negative, since they express costs per 1 ha of crops or per 1 animal. Nor may they have a too high positive value. This is the key problem. Negative regression coefficients are obtained for *other field crops*, *vegetables in covered areas*, *flowers and decorative plants*, *perennials*. The single very high value coefficient is for *flowers and decorative plants in covered areas*. The reason for that could be a small number of farms with these particular activities and small areas under these crops that make calculation of coefficients for the whole set of farms and all activities impossible – inexistence of common use of particular input and particular activity for all farms.

6) Then we can exclude from the further analysis the activities for which regression coefficients were negative or had unrealistic high positive values. This was done via selecting and discarding farms, which are influenced by the factors with negative or high positive value regression coefficients. The farms, cultivating *other field crops*, *vegetables in covered areas*, *flowers and decorative plants*, *perennials, flowers and decorative plants in covered areas* were discarded.

7) Then the regression analysis was repeated. If non-negative regression coefficients with reasonable standard errors were obtained, the problem has been successfully solved, we assumed.

We have to develop a procedure for the opposite outcome.

8) If possible, we may group the crops and animals with negative regression coefficients. For example, we may group grain with leguminous plants; vegetables with strawberries; (also we can try to put all cattle together, all pigs, sheep and goats). If grouping is possible, one is encouraged to do it.
If crops and animals cannot be grouped, we have to look at the volumes of production. If they are insignificant, we can discard farms with particular crops and animals. If this is not the case, discarding is not allowed and other approach should be used (it will be described later as *Application of "Activity-wise cost analysis")*. Anyway, we should note that discarding of farms with particular activities would lead to loosing the possibility to obtain cost allocation coefficients for these activities. A negative pesticide use coefficient was found for *oats*, therefore *oats* are grouped with *Other crops* and *Leguminous plants*. Regression analysis is done after a repeated grouping: the results are as follows (Table 4).

9) If it is impossible to obtain a non-negative coefficient for a crop or animal production activity, we could accept a little negative coefficients assuming that the analysed costs are not associated with the particular crop (and/or group of animals). For our purposes we assume these coefficients equal to 0.

10) We can also do regression analysis under the condition that intercept should not be a zero. The results are as in Table 2. To assess the credibility of the results, we calculated the sum of standard error squares of Regression coefficients (Table 4). The smaller the standard error, the smaller the sum of squares and the more credible the result. The credibility of the results is described also by determination coefficient R^2. The determination coefficient for pesticides is 0.74, thus, the validity of results is 74%.

Table 4. The first variant of pesticide regression coefficients

R^2=0,74	Intercept is 0		Intercept is not 0	
	Regression coefficients	Standard errors	Regression coefficients	Standard errors
Intercept	0,00	0,00	-246,08	131,66
Wheat	11,84	2,06	12,22	2,06
Rye	17,55	6,00	20,63	6,20
Barley	8,88	4,66	11,08	4,78
Other grain and leguminous plants	6,04	5,37	6,23	5,34
Potatoes	88,58	23,74	101,69	24,64
Sugar beet	125,74	7,00	127,25	7,02
Oil and fibre crops	39,36	9,93	42,33	10,01
Field vegetable	7,11	56,63	20,62	56,81
Vegetable in the specialised areas and strawberries	57,74	108,99	64,22	108,50
Fodder crops [in arable area]	2,31	2,22	2,01	2,21
Sum of the standard errors in square		**15892,10**		**33189,89**

11) Though it is not a must, we can try to do regression analysis for the other set as well. We can select the farms in which *Other field crops*, *field vegetables*,

vegetables in covered areas, vegetables in the specialised areas and strawberries, flowers and decorative plants, perennials, flowers and decorative plants in the covered areas are not raised and costs of pesticides are not 0. The results obtained for pesticides with the second variant are as in Table 5. The credibility rate, measured by sum of the standard errors in square, in this variant is higher, probably, because more activities with small production level are excluded from the regression analysis.

Table 5. The second variant of pesticides regression coefficients

R^2=0,80	Intercept is 0		Intercept is not 0	
	Regression coefficients	Standard errors	Regression coefficients	Standard errors
Intercept	0,00	0,00	-194,71	116,11
Wheat	11,75	1,78	11,85	1,78
Rye	10,25	4,53	12,10	4,65
Barley	8,13	3,75	9,99	3,90
Other grain and leguminous plants	6,31	4,39	6,20	4,38
Potatoes	61,46	28,28	76,89	29,64
Sugar beet	150,24	6,65	151,65	6,68
Oil and fibre crops	34,71	9,17	37,36	9,28
Fodder crops [in arable area]	3,14	1,59	3,08	1,58
Sum of the standard errors in square		***987,83***		***14552,02***

Seed costs

Under this exercise total seed costs (both farm produced and purchased) are analysed. Regression coefficients for seed may be obtained in 2 variants.

In the first variant, the farms were discarded, where other field plants, meadows and pastures, vegetables in covered areas, vegetables in specialised areas and strawberries, flowers and decorative plants, perennials, flowers and decorative plants in covered areas are raised and in which costs of seed are 0. (Table 6).

Table 6. The first variant of seed regression coefficients

R^2=0,77	Intercept is 0		Intercept is not 0	
	Regression coefficients	Standard errors	Regression coefficients	Standard errors
Intercept	0,00	0,00	40,76	609,62
Wheat	25,43	4,80	25,34	5,07
Rye	27,91	16,19	27,63	16,94
Barley	31,18	9,67	30,94	10,47
Other grain and leguminous plants	27,87	22,23	27,70	22,65
Potatoes	330,13	79,46	329,50	81,02
Sugar beet	76,17	14,57	76,03	14,91
Oil and fibre crops	37,57	28,01	37,06	29,38
Field vegetable	19,80	106,32	17,77	111,88
Sum of the standard errors in square		***19488,10***		***392735,67***

In the second variant, in addition to the first variant, farms with *field vegetables* were discarded. (Table 7).

Table 7. The second variant of the seed regression coefficients

R^2=0,90	Intercept is 0		Intercept is not 0	
	Regression coefficients	Standard errors	Regression coefficients	Standard errors
Intercept	0,00	0,00	354,88	88,69
Wheat	29,08	2,08	30,00	2,04
Rye	13,89	4,60	10,80	4,54
Barley	6,81	4,28	3,03	4,28
Other grain and leguminous plants	5,66	3,66	7,04	3,59
Potatoes	282,14	28,87	231,01	30,89
Sugar beet	89,67	5,78	86,20	5,69
Oil and fibre crops	18,21	9,69	12,87	9,53
Fodder crops [in the arable]	0,49	1,17	0,11	1,14
Meadows and pasturage	-2,92	0,92	-3,64	0,92
Sum of the standard errors in square		***1020,37***		***9001,00***

Fertiliser costs

Regression coefficients for fertiliser may be obtained in 2 variants.

a) In the first variant the value of manure produced in farm, valued in opportunity costs (for the calculation methodology - see Annex), was added to the costs of purchased fertiliser.
b) In the second variant only costs of purchased fertiliser are counted.

The both sub-variants (*Simple regression* and *Activity-wise cost analysis*) were used for fertiliser cost analysis.

Simple regression

In both variants the farms are selected, in which other field crops, oil and fibre crops, meadows and pastures, vegetables in covered areas, vegetable in specialised areas and

strawberries, flowers and decorative plants, perennials, flowers and decorative plants in covered areas are not grown and in which costs of fertiliser are not 0 (Table 8).

Table 8. The first variant of fertiliser regression coefficients

R^2=0,11	Intercept is 0		Intercept is not 0	
	Regression coefficients	Standard errors	Regression coefficients	Standard errors
Intercept	0,00	0,00	1810,53	1285,59
Wheat	41,03	22,97	42,42	22,96
Rye	49,57	63,30	37,83	63,75
Other grain and leguminous plants	21,19	30,15	13,40	30,61
Potatoes	48,33	359,28	-112,55	376,48
Sugar beet	247,75	73,33	235,20	73,76
Fodder crops in the arable	0,17	16,58	-2,02	16,63
Meadows and pasturage	9,20	12,37	5,96	12,57
Sum of the standard errors in square		***140332,94***		***1805877,29***

In the second variant the calculations only for purchased fertiliser (Table 9) are done.

Table 9. The second variant of fertiliser regression coefficients

R^2=0,86	Intercept is 0		Intercept is not 0	
	Regression coefficients	Standard errors	Regression coefficients	Standard errors
Intercept	0,00	0,00	-80,68	187,07
Wheat	47,45	3,41	47,40	3,42
Rye	48,69	9,34	49,18	9,42
Oats	8,78	10,26	9,30	10,35
Other grain and leguminous plants	-10,24	5,08	-9,93	5,14
Potatoes	157,06	52,47	164,29	55,14
Sugar beet	228,84	10,71	229,39	10,80
Fodder crops [in the arable]	-9,81	2,41	-9,71	2,42
Meadows and pasturage	-1,28	1,86	-1,14	1,89
Sum of standard errors in square		***3107,04***		***38397,02***

None of the two results are satisfactory. In the first case the *Sum of standard errors in square* is very high, whereas in the second case other grains but wheat, rye and oats, and also fodder crops have negative coefficients. Therefore the other sub-variant of regression analysis was applied.

"Activity-wise cost analysis" for the farm produced and purchased fertilisers

"Activity-wise cost analysis" is done according to the following procedure:

From the set of farms, for which the first variant of regression analysis was made, further the farms, in which *i* activity is observed (crop is grown) were picked-out. There are alternatives, for example- we can select only half of the farms, with the highest level of activity. Under *"Activity-wise cost analysis"* approach the regression analysis has to be repeated for every single activity. Regression analysis is done for these farms provided that regression intercept is 0, but only regression coefficient for *i* activity is taken as the

result. The rest of coefficients are ignored. Such regression coefficients were obtained for first variant of fertiliser calculation (Table 10).

Table 10. Regression coefficients of "Activity-wise *cost analysis"* for the farm produced and purchased fertiliser calculation

Plant	Regression coefficient	Standard error	Determination coefficient
Wheat	54,63	5,02	0,87
Rye	48,22	11,96	0,96
Other grain and leguminous plants	38,46	36,23	0,12
Potatoes	268,15	41,08	0,79
Sugar beet	143,60	59,90	0,96
Fodder crops [in the arable]	14,28	2,66	0,92
Meadows and pasturage	11,50	1,80	0,92

Veterinary costs

The following farm animal groups are selected for regression analysis:

- fattening cattle, including calves (<5 months), fattening young cattle (<1 year), bulls (1-2 years), fattening heifers (1-2 years), bulls (>2 years), cull dairy cows, other cows;
- breeding cattle, including breeding heifers (<1 years), breeding heifers (1-2 years);
- dairy cows;
- sows,
- other pigs, including piglets (<20 kg), growing sows (>50 kg), fattening pigs, other pigs;
- sheep and goats, including ewes and other sheep, goat breeding females, other goats,
- poultry, including laying hens, ducks, geese, chicken, other poultry;
- rabbits, including breeding females and other rabbits;
- horses, including working horses and horses for the selling;
- other animals, including bees.

At first regression analyses was applied for all the animal groups. However the obtained results had similar problems like in the case of pesticides. Therefore the next attempts were undertaken:

- in the first variant the farms with rabbits and bees were discarded. Results are presented in Table 11;
- in the second variant also the farms with *other animals*, *horses* were discarded. The results are presented in Table 12.

Table 11. The first variant of regression coefficients for the veterinary services

R^2=0,99	If intercept is 0		Intercept is not 0	
	Regression coefficients	Standard errors	Regression coefficients	Standard errors
Intercept	0,00	0,00	13,47	37,65
Dairy cows	15,16	1,57	15,04	1,61
Fattening cattle	2,22	0,70	2,22	0,71
Breeding cattle	2,05	2,03	2,14	2,04
Pigs	1,72	0,02	1,72	0,02
Sheep and goats	1,09	3,00	0,99	3,02
Poultry	0,11	0,00	0,11	0,00
Horses	4,89	6,67	4,73	6,70
Other animals	0,39	12,53	-0,43	12,75
Sum of standard errors in square		***217,47***		***1640,83***

Table 12. The second variant of regression coefficients for veterinary services

R^2=0,99	Intercept is 0		Intercept is not 0	
	Regression coefficients	Standard errors	Regression coefficients	Standard errors
Intercept	0,00	0,00	2,12	39,50
Fatting cattle	7,17	1,39	7,16	1,40
Growing cattle	0,50	2,18	0,51	2,20
Sheep and goats	0,08	6,29	0,05	6,33
Poultry	0,11	0,00	0,11	0,00
Dairy cows	14,63	1,76	14,61	1,80
Pigs	1,72	0,02	1,72	0,02
Sum of the standard errors in square		***49,37***		***1610,14***

Feed costs

Under this exercise total feed costs (both farm produced and purchased) are analysed. The same groups of cattle as in the previous analysis were used for the purpose of feed cost allocation. Hence the farms without *other animals*, *horses*, *rabbits* and *bees* have been selected (Table 13).

Table 13. Regression coefficients for feed costs

R^2=0,98	Intercept is 0		Intercept is not 0	
	Regression coefficients	Standard errors	Regression coefficients	Standard errors
Intercept	0,00	0,00	671,78	1653,35
Dairy cows	2,55	62,15	-1,04	62,85
Fattening and breeding cattle	307,18	56,48	307,61	56,55
Sows	117,21	72,91	112,96	73,74
Other pigs	54,61	3,86	54,80	3,89
Poultry	3,35	0,03	3,35	0,03
Sheep and goats	5,14	275,62	-4,51	276,96
Sum of standard errors in square in square		***88347,66***		***2822867,45***

Labour costs in agriculture

Simple regression analysis

At first, all crop activities (except fungus) and the animal groups analysed for the purposes of feed cost allocation, are used for regression analysis. Then the farms with *other field plants*, *oil and fibre crops*, *vegetable in the covered areas*, *vegetable in the specialised areas* and *strawberries*, *flowers and decorative plants*, *perennial plants*, *flowers and decorative plants in the covered areas*, *other animals*, *rabbits*, *horses* and *bees*, and in which registered labour consumption in agriculture is 0, were discarded. The results of regression are presented in Table 14.

Table 14. Regression coefficients for labour consumption

R^2=0,97	Intercept is 0		Intercept is not 0	
	Regression coefficients	Standard errors	Regression coefficients	Standard errors
Intercept	0,00	0,00	-2593,33	967,78
Wheat	42,26	16,75	37,55	16,61
Barley	136,58	37,69	166,18	38,78
Other grain and leguminous plants	24,55	26,06	27,40	25,73
Potatoes	435,18	261,10	640,76	268,71
Fodder crops [in the arable]	33,22	25,24	53,99	26,08
Meadows and pasturage	53,03	25,42	79,17	26,90
Sheep and goats	5,86	111,17	32,27	110,09
Poultry	0,83	0,01	0,83	0,01
Dairy cows	111,97	20,25	94,87	20,97
Pigs	13,61	0,34	13,69	0,34
Sum of the standard errors in square		***84604,97***		***1025197,30***

Activity-wise cost analysis

Because of high value of Sum of the standard errors in square, also the other sub-variant of regression analysis ("*Activity-wise cost analysis*") was applied. The results of that exercise are presented in Table 15. Also in this case the obtained results are hardly satisfactory, since several activities produced negative values.

Table 15. Regression coefficients for labour consumption from *Activity-wise cost analysis*

	Regression coefficient	Standard error	Determination coefficient
Wheat	-5,26	18,92	0,92
Barley	60,13	49,97	0,98
Other grain and leguminous plants	-56,42	25,71	0,92
Potatoes	195,01	226,62	0,92
Fodder crops [in the arable]	-11,54	26,38	0,89
Meadows and pastures	38,09	23,92	0,92
Sheep and goats	-89,06	0,00	1,00
Poultry	0,83	0,02	0,99
Cows	117,89	19,53	0,89
Pigs	13,73	0,13	0,99

2.3.4 Application of Two-stage regression analysis

This approach (described in 2.3.2. B) was used to allocate electricity and heating, fuel and lubricants, maintenance of machinery and buildings, rental costs and interest belong to the particular agricultural activities.

The following steps were undertaken.

The first stage: calculation of agriculture related part of respective costs:

- Farms, in which analysed costs are not 0, are selected. This has to be done to all cases of two-stages regression analysis.
- For further analysis farms with small land areas and small number of livestock are discarded. In all cases of two-stage regression analysis farms producing *other crops, vegetables in covered areas, vegetables in specialised areas and strawberries, flowers and decorative plants, perennials, flowers and decorative plants in covered areas, other farm animals, horses, rabbits, bees* are discarded. Also the farms with so-called *other agricultural earnings* was discarded, because the number of such farms was very small.
- We need to determine costs on 1 Ls of earnings in the crop production and livestock production, using regression analysis.
  This, so called first stage of the regression analysis, was done in 2 variants – where costs are the resultant (dependent) variable in both cases, whereas factorial variables are:
  1) in the 1st variant – produce value in crop production, livestock production, forestry, processing, other earnings;
  2) in the 2nd variant – aggregate agricultural produce value (crop production plus livestock production), forestry, processing, other earnings.

  The best of the result obtained could be identified either by the determination coefficient or by square of standard errors. The higher the determination coefficient, and the smaller the square of standard errors, the better the result.
  For fuel and lubricants, at the first stage of regression the following coefficients were obtained (Table 16.). Variant 1 seems as the best. The variant where agriculture is analysed by individual sectors – crop production and livestock production, and the intercept is 0 - gives the best results for the other costs as well.

Table 16. First stage regression for fuel and lubricants

	Intercept is 0				Intercept is not 0			
	1-st variant		2-nd variant		3-rd variant		4-th variant	
	Regression coefficients	Standard errors	Regression coefficients	Standard errors	Regression coefficients	Standard errors	Regression coefficients	Standard errors
Intercept					879,13	212,25	2396,14	271,55
Crop production	0,17	0,01			0,16	0,01		
Livestock production	0,02	0,00			0,02	0,00		
Total agriculture production			0,03	0,00			0,03	0,00
Forestry production	0,35	0,10	0,44	0,14	0,25	0,10	0,17	0,13
Processing of agricultural production	-0,13	0,02	-0,26	0,03	-0,12	0,02	-0,21	0,03
Other earnings	0,18	0,02	0,43	0,03	0,18	0,02	0,38	0,03
Sum of the standard errors in square		***0,01***		***0,02***		***45049,08***		***73741,20***
Determination coefficient		***0,76***		***0,49***		***0,76***		***0,56***

- From the total farm costs we need to subtract costs linked to forestry, processing and other earnings. The above costs are obtained by multiplying the obtained regression coefficients with value of forestry, processing and other earnings. Thus we arrive at respective costs in agriculture.

Second stage of regression is done in a similar manner as the one-stage regression analysis. Resultant feature – respective costs in agriculture; factorial features – areas planted with crops and number of various farm animals.

Further we need make the simple regression analysis for the accordant costs in the agriculture in dependence on plant areas and cattle number. Following result was got for fuel and lubricants (Table 17), for electricity and heating costs (Table 18), for maintenance for machinery and buildings (Table 19), paid interest (Table 20).

Table 17. Regression coefficients for fuel and lubricant costs

R^2=0,85	Intercept is 0		Intercept is not 0	
	Regression coefficients	Standard errors	Regression coefficients	Standard errors
Intercept	0,00	0,00	732,98	173,39
Wheat	15,07	2,74	15,37	2,65
Rye	45,57	7,05	42,02	6,87
Barley	14,89	6,57	4,95	6,78
Oats	21,28	8,51	19,27	8,25
Other grain and leguminous plants	34,41	10,75	43,32	10,62
Potatoes	47,49	48,79	-27,21	50,43
Sugar beet	34,42	10,30	29,34	10,04
Oil and fibre crops	13,40	13,99	3,72	13,74
Fodder crops [in the arable]	7,66	5,59	2,65	5,54
Meadows and pasturage	-1,88	4,92	-8,31	5,00
Sheep and goats	8,43	20,05	0,62	19,50
Poultry	0,05	0,00	0,05	0,00
Fatting cattle	23,40	8,02	30,50	7,94
Growing cattle	2,18	9,32	8,47	9,15
Cows	7,78	10,97	9,48	10,63
Pigs	0,71	0,06	0,69	0,06
Sum of the standard errors in square		***3699,64***		***33875,33***

Table 18. Regression coefficients for electricity and heating costs

R^2=1	Intercept is 0		Intercept is not 0	
	Regression coefficients	Standard errors	Regression coefficients	Standard errors
Intercept	0,00	0,00	141,89	120,37
Grain and leguminous plants	-1,40	1,22	0,98	1,26
Potatoes	1,99	28,51	-13,09	29,94
Sugar beet	5,72	6,99	-6,37	6,99
Oil and fibre crops	1,90	11,01	-3,11	11,03
Fodder crops [in the arable]	-3,92	3,70	3,15	3,74
Meadows and pasturage	-9,37	3,17	8,18	3,31
Fatting cattle	12,75	4,76	-11,03	4,95
Growing cattle	13,03	6,04	-11,52	6,15
Sheep and goats	2,99	13,24	-4,40	13,25
Poultry	0,45	0,00	-0,45	0,00
Dairy cows	24,97	7,14	-25,08	7,12
Pigs	5,25	0,04	-5,24	0,04
Sum of the standard errors in square		***1293,53***		***15871,20***

Table 19. Regression coefficients for machinery and buildings maintenance costs

R^2=0,81	Intercept is 0		Intercept is not 0	
	Regression coefficients	Standard errors	Regression coefficients	Standard errors
Intercept	0,00	0,00	445,64	161,77
Grain and leguminous plants	7,90	1,58	6,61	1,63
Potatoes	111,13	39,93	79,84	41,00
Sugar beet	41,53	9,39	39,24	9,31
Oil and fibre crops	19,58	13,04	14,72	12,98
Fodder crops [in the arable]	-4,38	5,15	-7,00	5,17
Meadows and pasturage	-9,09	4,49	-13,32	4,69
Fatting cattle	17,71	6,85	23,51	7,08
Growing cattle	11,60	9,15	17,20	9,25
Poultry	0,04	0,02	0,04	0,02
Sheep and goats	5,29	18,84	0,93	18,66
Cows	15,01	10,37	14,11	10,24
Pigs	1,65	0,06	1,63	0,06
Sum of the standard errors in square		***2495,35***		***28746,59***

Table 20. Regression coefficients for the interest paid

R^2=0,92	Intercept is 0		Intercept nav0	
	Regression coefficients	Standard errors	Regression coefficients	Standard errors
Intercept	0,00	0,00	1224,40	528,04
Grain and leguminous plants	7,26	3,95	3,75	4,17
Potatoes	237,39	102,99	175,40	104,94
Sugar beet	233,76	24,36	228,67	24,10
Oil and fibre crops	50,92	33,89	38,05	33,85
Fodder crops [in the arable]	3,03	12,28	-5,27	12,62
Meadows and pasturage	-12,74	13,02	-22,24	13,47
Sheep and goats	5,53	46,25	-6,15	45,85
Poultry	0,04	0,04	0,03	0,04
Cows	12,60	9,77	18,37	9,95
Pigs	5,63	0,14	5,59	0,14
Sum of the standard errors in square		***14918,73***		***294123,59***

Land rent does not depend on number of animals, so we discard factorial variables about animal number (Table 21).

Table 21. Regression coefficients of land rent

R^2=0,57	Intercept is 0		Ja Intercept nav0	
	Regression coefficients	Standard errors	Regression coefficients	Standard errors
Intercept	0,00	0,00	-130,39	227,68
Grain and leguminous plants	4,69	1,40	4,87	1,44
Potatoes	53,75	36,05	57,66	36,94
Sugar beet	12,57	9,47	13,98	9,85
Oil and fibre crops	3,47	23,58	6,73	24,42
Fodder crops [in the arable]	1,84	1,55	1,84	1,56
Meadows and pasturage	5,61	4,64	6,11	4,76
Sum of the standard errors in square		***1971,44***		***53925,05***

3. COMPARISON OF CALCULATED AND CURRENTLY USED INPUT USE COEFFICIENTS AND CONSISTENCY TO EAA DATA

The comparison of calculated and currently used input use coefficients per activity is presented in the next tables. An adjustment coefficient is shown in the last row of every table.

This coefficient was calculated individually for each type of input. Its purpose is to align the input part with the use part. The input costs per activity were multiplied with the activity level and the sum was compared with the total value of that input in the EAA. The used amount is equalled to inputs, and the coefficient obtained is applied on costs for each individual production activity

3.1. Comparison of new and old methodology

Table 22. Input coefficients for pesticides, seeds and fertilisers

Production activity	Pesticides		Seeds		Fertilisers		Land rent	
	New approach	Old approach	New approach	Old approach	New approach	Old approach	New approach	Old approach
Winter wheat	11.75	25.00	29.08	15.99	54.63	41.25	4.69	3.88
Spring wheat	11.75	31.20	29.08	15.99	54.63	23.98	4.69	3.88
Rye	10.25	13.00	13.89	12.78	48.22	35.73	4.69	3.88
Barley	8.13	13.00	6.81	12.93			4.69	3.88
Other grain products	6.31	13.00	5.66	12.15	38.46	36.43	4.69	3.88
Pulses	6.31	34.00	5.66	12.51	38.46	18.39	4.69	3.88
Potatoes	61.46	28.27	282.14	156.37	268.15	83.38	53.75	3.88
Sugar beets	150.24	102.00	89.67	65.00	143.60	66.33	12.57	3.88
Rape	34.71	39.00	18.21	10.14			3.47	3.88
Flax	34.71	45.00	18.21	4.89			3.47	3.10
Fodder roots	3.14	8.50	0.49	65.00	14.28	53.70	1.84	0.59
Maize	3.14	6.50	0.49	24.00	14.28	46.38	1.84	0.59
Meadows, pastures				0.08	11.50		5.61	0.59
Adjustment coefficient	0.372	0.329	0.700	0.762	0.484	0.969	0.141	0.513

Table 23. Input coefficients for veterinary expenses and feeding stuffs

Production activity	Veterinary expenses		Feeding stuffs	
	New approach	Old approach	New approach	Old approach
Fattening cattle	2.22	1.20	307.18	86.99
Breeding cattle	2.05	1.20	307.18	86.99
Sheep and goats	1.09	1.03	5.14	24.62
Poultry	0.11	0.135	3.35	1.642
Dairy cow	15.16	10.64	2.55	147.19
Pigs	1.72	0.85	54.61	69.13
Sow			117.21	156.04
Adjustment coefficient	0.923	0.932	0.934	0.915

Table 24. Input coefficients for petrol and lubricants, electricity and fuel, maintenance of materials and buildings, interest paid

Production activity	Petrol, lubricants		Electricity, fuel		Maintenance of materials, buildings		Interest paid	
	New approach	Old approach	New approach	Old approach	New approach	Old approach	New approach	Old approach
Winter wheat	15.07	16.46		1.35	7.90	12.67	3.75	34.84
Spring wheat	15.07	14.90		1.35	7.90	11.46	3.75	31.53
Rye	45.57	14.90		1.28	7.90	11.46	3.75	31.53
Barley	14.89	15.38		1.28	7.90	11.84	3.75	31.53
Oats	21.28	16.37		1.16	7.90	11.84	3.75	32.55
Other grain products	34.41	16.37		1.28	7.90	12.67	3.75	12.67
Pulses	34.41	17.54		0.97	7.90	13.50	3.75	37.11
Potatoes	47.49	31.56	1.99	1.35	111.13	24.29	175.40	66.79
Sugar beets	34.42	36.81	5.72	0.70	41.53	28.32	228.67	77.89
Rape	13.40	19.47	1.90	0.43	19.58	14.98	38.05	41.19
Flax	13.40	25.99	1.90	0.55	19.58	15.77	38.05	43.36
Fodder roots	7.66	21.42		0.70	0	28.32		6.74
Maize	7.66	10.15		0.55	0	6.40		21.48
Fattening cattle	23.40	1.17	12.75	3.46	17.71	3.20		3.42
Breeding cattle	2.18		13.03		11.60			3.42
Dairy cows	7.78	1.90	24.97	10.00	15.01	1.65	18.37	13.17
Sheep and goats	8.43	0.36	2.995	1.14	5.29	0.76	5.53	1.96
Poultry	0.05	0.012	0.45.	0.42	0.04	0.03	0.03	0
Pigs	0.71	0.83	5.25	2.05	1.65	0.99	5.59	1.96
Sow	0.71	1.43	5.25	4.09	1.65	1.99		
Adjustment coefficient	0.832	1.281	0.535	0.868	0.654	0.335	0.189	0.125

4. SUMMARY AND CONCLUSIONS

Eurostat currently is setting up the so-called Agricultural Information System (AgrIS). The major objective of AgrIS is to restructure NewCRONOS annual agricultural statistics into a coherent information system and to make them more easily comprehensible, usable and even improvable by all the users concerned.

Furthermore, AgrIS will provide the data - basis for Agricultural Sector Modelling. The most difficult part in this work is to obtain the data about the use of intermediate inputs per activity.

The objective of this study was to develop and test an approach for using data from the FADN and other data sources for generating input coefficients per Agricultural Activity. The focus of the study is a methodology, which could be applied on a regular basis and which would provide a regular data flow to Eurostat concerning input coefficients.

In the perspective, FADN might become as the main information source in respect to intermediate consumption in agriculture, where, applying the calculated coefficients to crop areas or to number of animals, one can arrive at the respective aggregate intermediate consumption item, to be used for EAA.

This particular study could be counted as the first stage in the solving of the problem, and its particular task is to try to meet Eurostat minimum requirements – to obtain data on the key cost items by Agricultural Activity. Apart from that, there was an attempt to break down the labour costs, rent and interest payments as well.

The first step of the project was to examine the current situation in Latvia regarding the existence of cost items by Agricultural Activity, and their data source.

Since 1995, EXCEL worksheets have been used for calculations of EAA in Latvia, using ABTA approach used by SPELL. This means that input use coefficients by Agricultural Activity are already available. The main data source for obtaining these coefficients are LAAC gross margin calculations, small surveys of farms, normative data (obtained via research), as well as expert estimations. The first chapter of the report presents data source in more detailed way.

Since the main task of the project was to use FADN data in order to obtain input use coefficients by Agricultural Activity, FADN database was analysed, including its compliance with the present EAA data set and AgrIS requirements. The analysis method to be used for the study purpose was selected.

Input use items required by AgrIS information system were compared with the items in the existing Latvian ABTA table and those available in FADN data base. Information, needed to obtain data on input use items at AgrIS top aggregation level was available both in ABTA table and in FADN database. Similarly, Agricultural Activities were compared and assessed. FADN database has a very detailed breakdown by Agricultural Activity. However, for each individual activity the number of farms varied, and in many cases it was too small. Therefore, for the purpose of the study, an Agricultural Activity aggregation level was selected which allowed obtaining a data set, which could meet the representation requirements for mathematical analysis. The selected aggregation level allows comparing FADN data with ABTA table data; however, it does not fully correspond to the minimum list of activities, specified in AgrIS system.

To calculate input use coefficients, a multiple linear regression (hereafter – regression) method was selected. This choice was based on our own considerations, as well as on the experience of other countries (namely, Denmark), and recommendations (distributed by ASA). Following the theory, this method is applied to analyse relationship between several independent variables and one dependent variable. In case of our problem, crop areas and number of animals were taken as independent variables, while the dependent variable was common individual cost items, which, of course, depend on the number of ha or on number of animals. The obtained regression coefficients basically describe the effect of each independent variable (area or number of animals) on the dependent variable (total respective costs). In our case, regression coefficients would be the respective costs by Agricultural Activity per ha or per animal.

As in input item, FADN data set was prepared for each individual cost item to apply regression method, selecting farms where respective costs are relevant. For example, some of the cost items are relevant exclusively to crop farming or to animal production – pesticides apply exclusively to crop production. In this case farms with dominating livestock production were excluded from the input data set for analysis.

The initial analysis for individual production activities ended up with negative regression coefficients, which are not feasible. Assessment of FADN data led to a conclusion that negative figures are generated by production activities, which have comparatively small scale (small areas or small number of animals) or represented in very few farms. Therefore, activities with possibly similar cost items were grouped. Some farms with very specific production activities (e.g. fungus) were altogether excluded from the input data set.

Some cost items, for example, electricity and fuel for heating, motor fuel, maintenance costs are applicable to entire agriculture and also to non-agricultural secondary activities. Second step was to break down agriculture-related costs by Agricultural Activity.

The first attempts to break down labour costs and fertiliser costs by Agricultural Activity did not result in credible coefficients. They were either negative, or else, very large positive figures. In these cases additionally the so-called "*Activity-wise cost analysis*" method was applied, however, even this did not end up in credible results in all cases. This means that in the analysis in question, the selected method could not find a constructive (credible) solution for the specified data set.

Multiple linear regression equation includes also intercept. If its value is not a zero, it means that all costs have not been distributed by Agricultural Activity or vice versa - over distributed. The study included calculations both for the cases when intercept is a 0 and when it is not, and there were attempts to find the most credible solution.

Credibility of the results and their quality can be valuated also by such criteria as standard error and determination coefficient. The result is better if the coefficient standard errors are smaller and determination coefficient is larger.

Applying regression analysis to variously grouped FADN data and applying different assumptions, varying results were obtained, of which the most credible were selected for further analysis. Yet, due to various circumstances and some issues that remain unsolved so far, even these results by no means can be taken as final and correct.

The data quality is one of the prerequisites for obtaining credible coefficients. Latvian FADN database contains information on different farms – small farms with a couple of ha of agricultural land and a few animals, as well as large farms with 100 and over ha of land and a large number of animals. Consequently, technologies and even farming strategies

may largely vary between the farms. The farms also differ by type of farming. There are mixed farms where it is hard to identify the core business, and there are market-oriented farms with comparatively narrow specialisation.

The quality of data recorded at the farms and entered into FADN data base might also be the cause of the problem, especially the stock and output evaluation practices and the scales used for that, the record accuracy of use of labour and other inputs; differing prices of the used resources.

These could be the reasons for obtaining negative or large positive and incredible input use coefficients. To improve the results, the farms were grouped in various ways by activities, and, on separate occasions, were altogether excluded from input data set. However, it is impossible to exclude too many farms, because this reduces the representative quality of the obtained set.

Similarly, it is premature to conclude that the obtained coefficients might describe the average level of agriculture in Latvia, since the Latvian FADN database is under development and does not represent the entire agricultural sector that will be possible only after carrying out agricultural census in Latvia.

The technical calculations were done applying regression analysis tool, included in EXCEL. Yet, it does not allow to define with sufficient accuracy the limiting factors for the problem solution, which, possibly, could solve several issues related to result credibility and logic – negative figures and extreme figures. Similarly, it allows a limited number of variables, which does not allow including simultaneously all activities that are analysed. In future it is recommendable to use statistics software, e.g. SPSS or SAS, which offers more ample possibilities to formulate the problem.

The key intermediate consumption and other cost item breakdown coefficients by Agricultural Activity can be counted as one of the result of the study.

The input use coefficients obtained in the last phase of the study were compared with the existing ones. Presently it would not be correct to evaluate whether the obtained coefficients are better, because only the first step has been taken to develop a new methodology, and the method has not been sufficiently elaborated. Likewise, the analysed data set itself calls for further development and improvement, too.

In future the use of FADN data base for the purpose of breaking down the costs by Agricultural Activity could be a good solution, because FADN data base contains information on farms which do accounting, thus, actual costs are available. The currently used data sources are more built on normative acts and rule of thumb rather than actual average costs in Latvian farms per Agricultural Activity unit.

Prior to be able to use FADN data, there is much work to be done which might take a couple of years, and the first steps could be:

- Improvement of FADN data set- data quality and representativeness - as the prerequisite for this is a new farm sample designed on the basis of agricultural census data.
- A thorough input data preparation for the purpose of each cost item breakdown by Agricultural Activity;
- Regression analysis technology should be improved – documentation of analysis scenario, building an interface, result export etc.;

- Comparing the results obtained with the above described method with other calculation methods, for example, the one used in INRA (France) and recommended for use in the EU member countries.

For the time being, to obtain cost breakdown by Agricultural Activity, also alternative data sources and calculation methods should be used in Latvia.

ANNEX

Calculation of total costs for fertilising

There are total costs of fertilisers in EAA, but there are farms with different level of intensity of animal breeding, some farms use only purchased fertilizers, some farms can use manure instead of buying fertilisers, but some farms have a problem to utilize a manure. Every group of animals gives different amount of manure per year and the content of these manure are different. And if there are data about yield of manure from one animal per year and about content of manure, then we can calculate the utility of manure. We have used standards of manure in Latvia, which were worked out in 1999.

Table 25. Temporary Standards of Manure in Latvia

System of animal keeping	Kind of manure	Yield per year, t	Content, kg/t naturally raw manure		
			N	P	K
Sow with 18n piglet, till piglet attain weight 20 kg					
Continuous floor	Litter manure	4.6	4.6	3.5	3.6
Feedlot pig, live weight 20-100 kg					
Continuous floor	Liquid fertilisers	3.6	3.3	1.6	2.8
Dairy cow, milk yield 3500- 5000 kg per year					
Leashed. Continuous floor	Liquid fertilisers	22	2.3	0.9	1.9
Calf (neat), till 6 month old					
Leashed. Continuous floor	Liquid fertilisers	6.0	1.9	0.8	1.5
Heifer, 6 to 24 month old					
Leashed. Continuous floor	Liquid fertilisers	15	2.2	0.9	1.8
Cattle, from 6 month old till 450 kg live weight (26 month)					
Leashed. Continuous floor	Liquid fertilisers	20.5	2.2	0.9	1.8
Horse					
Continuous floor	Litter manure	8	5.2	3.6	7.5
Sheep					
Deep byre	Litter manure	0.9	7.8	4.7	10.5
Hen					
Continuous floor	Liquid fertilisers	0.1	10.6	7.2	3.8
Source: Conditions for good practice of agriculture in Latvia, LLU, Jelgava 1999					

In calculation of total costs for fertilising we can use following formula:

1) K*0.7=R;

K – yield of manure from one animal per year;

R - manure for enforce;

0.7 – by opinion of scientists 30 % of manure lapse in time of storage.

2) R*v (N, kg)=N^1 ,

R*v (P, kg)= P^1,

R*v (K kg)= K^1,

N^1 - amount of N (kg) from manure of one animal per year;

P^1 - amount of P (kg) from manure of one animal per year;

K^1 - amount of K (kg) from manure of one animal per year;

3) N^1*p(N)+ P*p(P)+ K*p(K)= S

p(N) – price of 1 kg N;

p(P) - price of 1 kg P;

p(K) - price of 1 kg K;

S – value of farm produced manure, measured as opportunity costs (value of possibly replaced chemical fertiliser)

Table 26. Prices of N; P; K in Latvia (1999)

Kind of fertilisers	Amount, kg	Price for 1 kg of fertilisers, LVL
N	1	0.29
P	1	0.46
K	1	0.11

Table 27. Results of Calculation

Animals	K	v (N;P;K), kg			R	N^1	P^1	K^1	N^1;P^1;K^1*price (N;P;K)			S
	Yield of manure from 1 animal per year, t	Content, kg/t naturally raw manure			K*0.7, t	R*v (N;P;K) kg			N^1*p(N)	P*p(P)	K*p(K)	
		N	P	K		N^1	P^1	K^1	0.29 LVL/ kg	0.46 LVL/ kg	0.11 LVL/ kg	
Sow	**4.6**	4.6	3.5	3.6	3.22	14.81	11.27	11.59	4.29	5.18	1.27	9.48
Fatten. pigs	**3.6**	3.3	1.6	2.8	2.52	8.32	4.03	7.06	2.41	1.85	0.78	5.47
Dairy cows	**22**	2.3	0.9	1.9	15.4	35.42	13.86	29.26	10.27	6.38	3.22	16.65
Calf	**6**	1.9	0.8	1.5	4.2	7.98	3.36	6.30	2.31	1.55	0.69	5.03
Heifer	**15**	2.2	0.9	1.8	10.5	23.1	9.45	18.9	6.69	4.35	2.08	13.89
Cattle	**20.5**	2.2	0.9	1.8	14.35	31.57	12.92	25.83	9.16	5.94	2.84	17.94
Horse	**8**	5.2	3.6	7.5	5.6	29.12	20.16	42	8.44	9.27	4.62	17.71
Sheep	**0.9**	7.8	4.7	10.5	**0.63**	**4.91**	**2.96**	**6.62**	**1.43**	**1.36**	**0.73**	**2.79**
Hen	**0.1**	10.6	7.2	3.8	**0.07**	**0.74**	**0.50**	**0.27**	**0.22**	**0.23**	**0.03**	**0.45**