

Modern Classification: Principles and Documentation - LAND

Authors: Kessler G., Markevich A., Arkhina A.

INTRODUCTION

Historical statistics are coined in the categories and terms of the period – what we refer to within the framework of the Electronic Repository of Russian Historical Statistics as ‘historical classifications’. These categories can even differ within one and the same period, due to the existence of different spelling variants and synonyms. Particularly for the earlier benchmark-years 1795 and 1858 this is often the case. Standardising these categories across data-sets in order to maximise the comparability of the data is one of the central aims of the Electronic Repository of Russian Historical Statistics (ERRHS).

To achieve this, we have:

(a) unified and systematised the 'historical' categories used in the sources and;

(b) matched and marked them accordingly to categories used in accepted modern, and highly standardised systems of classification (NACE, CLEMS, PST and others).

in which:

(a) The standardisation of historical categories into what we will refer to as ‘unified historical classifications’ serves the purpose of making data comparable within one and the same period and facilitates querying the data.

(b) the reliance on modern classifications to tag the data resolves the issue of comparability over time. Unifying 'historical classifications' does not render data comparable between cross-sections. Different grouping principles might have been used, and the meaning of terms and categories apparently similar, can in fact have changed over the years. Coding the data for different years using one and the same classification scheme does, however, solve this problem.

On the Ristat-portal users can query the database using the historical or the modern classification, and aggregate the data at different levels of hierarchy. This query tool is accessed from the left-hand menu at ‘STATISTICS/TOPICS’. A radio-button allows the user to toggle between historical and modern classification, changing the set of categories available for selection under ‘4. Choose indicators’.

Full disaggregated data-sets in database format (one row=one record) with simultaneous historical and modern classification can be downloaded from the left-hand menu at ‘STATISTICS/FILE CATALOGUE’. The historical classification is in the fields ‘HistClass1’, ‘HistClass2’.....’HistClass10’. The modern classification is in the fields ‘Class1’, ‘Class2’.....’Class10’. Below, we expand on the methodology used in applying modern classifications.

GENERAL PRINCIPLES OF THE APPLICATION OF MODERN CLASSIFICATIONS

Modern classifications are, as a rule, hierarchical, and consist of nested categories, which allow for the classification of each and every historical variable, even if at different levels of detail. Branches and sub-branches of industry can be taken as a good example. In case a certain figure cannot be matched to a single sub-branch of industrial production, it might well be feasible to determine the more general branch of production that it relates to, one level up in the hierarchy.

However, such a procedure rarely allows one to reconstruct all categories of the modern classification, because the necessary level of detail of the data to do so might simply not be available in the source. Only if each historical category matches to only one category in the modern classification (a relationship of many-to-one), and all necessary historical categories are identifiable and available in the

source ('functional completeness') can we speak of a full reconstruction of a category from the modern classification. This is usually the case only at the higher, as well as at the lower levels of aggregation, whereas at intermediate levels the mismatches are a more vexing problem.

This notwithstanding, we have categorised all data in the database in terms of a modern classification, either through full reconstructions, or by distributing the available historical categories among modern 'baskets'. In the next sections, we will specifically elaborate on our methodology with regards to the modern classification of ERRHS' data on land.

7.01 - LAND BY LAND USE

There are two widely accepted systems of classification relating to land: LUCAS, used by Eurostat, and the classification of Land Inputs used at the Food and Agriculture Organization of the United Nations (FAO).

The LUCAS system used by Eurostat is a very sophisticated tool, which, considering the type of data we need to classify, has one advantage and one disadvantage.

The advantage is that it contains elaborate subcategories allowing for a highly differentiated classification of arable land according to crops grown or types of animal husbandry practiced.

The disadvantage is that it consists of separate classifications according to land cover and land use, the first relating to natural phenomena (forest, water, arable etc.), and the second to the socio-economic use made of the land. Only the latter category includes a graph for "non-arable land", a category which figures prominently in most of our data, but does not contain categories for water, forest etc, as these do not relate to land use, but to land cover.

Weighing these advantages and disadvantages we have opted for the FAO classification of Land Inputs, because it fits our data better. We operate a subset of the most relevant overarching categories from the classification, which is appended to this methodological paper (cf. Appendix nr. 1).

Users interested in the data on crops or types of animal husbandry practiced, are advised to refer to the data in their historical classification. If need be, they could use the elaborate FAO classification of crops and livestock products or the NACE classification of types of economic activity to add this dimension to the data.

7.02 - Land by land tenure

The Electronic Repository offers data on land tenure for 1795, 1858, and 2002. Applying a uniform modern classification to these data is complicated by two circumstances: (1) the strong dependence of land tenure forms on the legal context of a particular period and (2) the lack of a generally accepted classification of forms of land tenure. We assume that the second circumstance follows from the first - the specificity of place and time tends to make a general classification meaningless. In view of this, we decided not to apply a modern classification to the data in this section.

7.03 - Land prices

Data on land prices are contained in the resource only for 1897, expressed in rubles per unit of area measurement. No modern classification is applied to these data.

7.04 - Land lease prices

The data on the price of land rent is contained in the resource only for 1897, expressed in rubles per unit of area measurement. No modern classification is applied to these data.

...

December 2021

APPENDIX 1 - SUBSET OF CLASSIFICATION LAND INPUTS FAO, AS USED IN THE ELECTRONIC REPOSITORY OF RUSSIAN HISTORICAL STATISTICS

Author: Gijss Kessler

Source: Food and Agriculture Organization of the United Nations, *Data Structure, Concepts and Definitions common to FAOSTAT and CountrySTAT framework*, June 2010,

http://countrystat.org/resources/metadata_en.pdf [last accessed on 8 December 2021]

Hierarchy:

6600 - Country area	6680 - Inland water				
	6601 - Land area	6670 - Other land			
		6661 - Forest	6714 - Primary forest		
			6716 - Planted forest		
			6717 - Other naturally regenerated forest		
			6630 - Temporary crops		
		6610 - Agricultural area	6620 - Arable land and Permanent crops	6621 - Arable land	6633 - Temporary meadows and pastures
					6640 - Fallow land (temporary)
					6650 - Permanent crops
			6655 - Permanent meadows and pastures	6656 - Perm. meadows & pastures - Cultivated	
				6659 - Perm. meadows & pastures - Nat. growing	

Definitions:

Item Code	Item Name	Definition
6600	Country area	Country area, area of the country including area under inland water bodies, but excluding offshore territorial waters.
6601	Land area	Land area is the total area of the country excluding area under inland water.
6610	Agricultural area	Agricultural area, this category is the sum of areas under "Arable land", "Permanent crops" and "Permanent pastures".
6620	Arable land and Permanent crops	Arable land and Permanent crops, this category is the sum of areas under "Arable land" and "Permanent crops".
6621	Arable land	Arable land is the land under temporary agricultural crops (multiple-cropped areas are counted only once), temporary meadows for mowing or pasture, land under market and kitchen gardens and land temporarily fallow (less than five years). The abandoned land resulting from shifting cultivation is not included in this category. Data for "Arable land" are not meant to indicate the amount of land that is potentially cultivable.
6630	Temporary crops	Temporary crops is all land used for crops with a less than one-year growing cycle and which must be newly sown or planted for further production after the harvest.
6633	Temporary meadows and pastures	Temporary meadows and pastures is the land temporarily cultivated with herbaceous forage crops for mowing or pasture. A period of less than five years is used to differentiate between temporary and permanent meadows.
6640	Fallow land (temporary)	Fallow land (temporary) is the arable land that is not seeded for one or more growing seasons. Land is not considered temporarily fallow unless it has been, or is expected to be, kept at rest for at least one agricultural year. Fallow land temporarily used for grazing should be classified as "fallow" if the land is normally used for growing temporary crops. The maximum idle period is usually five years; land remaining fallow longer may acquire characteristics requiring it to be reclassified, such as "permanent meadows and pastures" (if used for grazing), "forest or other wooded land" (if overgrown with trees), or "other land".
6650	Permanent crops	Permanent crops is the land cultivated with long-term crops which do not have to be replanted for several years (such as cocoa and coffee); land under trees and shrubs producing flowers, such as roses and jasmine; and nurseries (except those for forest trees, which should be classified under "forest"). Permanent meadows and pastures are excluded from land under permanent crops.
6655	Permanent meadows and pastures	Permanent meadows and pastures is the land used permanently (for a period of five years or more) for herbaceous forage crops, either cultivated or naturally growing. A period of five years or more is used to differentiate between permanent and temporary meadows.
6656	Perm. meadows & pastures - Cultivated	Permanent meadows and pastures is the land used permanently (for a period of five years or more) for herbaceous forage crops that are managed and cultivated.
6659	Perm. meadows & pastures - Nat. growing	Permanent meadows and pastures - Naturally growing is the land used permanently (for a period of five years or more) for herbaceous forage crops that is naturally growing.

6661	Forest	<p>Forest area is the land spanning more than 0.5 hectares with trees higher than 5 metres and a canopy cover of more than 10 percent, or trees able to reach these thresholds in situ. It does not include land that is predominantly under agricultural or urban land use. Forest is determined both by the presence of trees and the absence of other predominant land uses. The trees should be able to reach a minimum height of 5 metres (m) in situ. Areas under reforestation that have not yet reached but are expected to reach a canopy cover of 10 percent and a tree height of 5 m are included, as are temporarily unstocked areas, resulting from human intervention or natural causes, which are expected to regenerate. Includes: areas with bamboo and palms provided that height and canopy cover criteria are met; forest roads, firebreaks and other small open areas; forest in national parks, nature reserves and other protected areas such as those of specific scientific, historical, cultural or spiritual interest; windbreaks, shelterbelts and corridors of trees with an area of more than 0.5 ha and width of more than 20 m; plantations primarily used for forestry or protective purposes, such as: rubber-wood plantations and cork, oak stands. Excludes: tree stands in agricultural production systems, for example in fruit plantations and agroforestry systems. The term also excludes trees in urban parks and gardens. Forest is composed of the following sub-categories: "Primary forest", "Other naturally regenerated forest" and "Planted forest".</p>
6670	Other land	<p>Other land is the land not classified as Agricultural land and Forest area. It includes built-up and related land, barren land, other wooded land, etc.</p>
6680	Inland water	<p>Inland water is the area occupied by major rivers, lakes and reservoirs.</p>
6714	Primary forest	<p>Naturally regenerated forest of native species, where there are no clearly visible indications of human activities and the ecological processes are not significantly disturbed. Explanatory note: 1. Some key characteristics of primary forests are: • they show natural forest dynamics, such as natural tree species composition, occurrence of dead wood, natural age structure and natural regeneration processes; • the area is large enough to maintain its natural characteristics; • there has been no known significant human intervention or the last significant human intervention was long enough ago to have allowed the natural species composition and processes to have become re-established.</p>
6716	Planted forest	<p>Forest predominantly composed of trees established through planting and/or deliberate seeding. Explanatory notes 1. In this context, predominantly means that the planted/seeded trees are expected to constitute more than 50 percent of the growing stock at maturity. 2. Includes coppice from trees that were originally planted or seeded. 3. Includes rubberwood, cork oak and Christmas tree plantations. 4. Excludes self-sown trees of introduced species</p>
6717	Other naturally regenerated forest	<p>Naturally regenerated forest where there are clearly visible indications of human activities. Explanatory notes: 1. Includes selectively logged-over areas, areas regenerating following agricultural land use, areas recovering from human-induced fires, etc. 2. Includes forests where it is not possible to distinguish whether planted or naturally regenerated. 3. Includes forests with a mix of naturally regenerated trees and planted/seeded trees, and where the naturally regenerated trees are expected to constitute more than 50 percent of the growing stock at stand maturity.</p>

Appendix 2 - User license and correct reference to the data from the Electronic Repository of Russian Historical Statistics

The data available from the Electronic Repository of Russian Historical Statistics are licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

This license allows you to share, edit and re-use the data from the Electronic Repository of Russian Historical Statistics for non-commercial purposes, provided that due reference is made to the Electronic Repository of Russian Historical Statistics and derivative products or materials are distributed under the same license.

By downloading and using data from the Electronic Repository of Russian Historical Statistics the user agrees to the terms of this license.

For an overview of the main terms of the license, cf.:

<http://creativecommons.org/licenses/by-nc-sa/4.0/>

For the full text of the license, cf.:

<http://creativecommons.org/licenses/by-nc-sa/4.0/legalcode>

Please, note, that the Electronic Repository of Russian Historical Statistics is not a source publication. The exact source for each figure is listed, but the database in its entirety is a complex product of dedicated research, analysis and interpretation. Therefore, please adhere to the same standards of citation that apply to other products of research like books and articles and refer to the source of your data as listed below:

Kessler, Gijs and Andrei Markevich, Electronic Repository of Russian Historical Statistics, 18th - 21st centuries, <https://ristat.org/>, Version I (2021): [Datatype number and name] [benchmark-year].

Example:

Kessler, Gijs and Andrei Markevich, Electronic Repository of Russian Historical Statistics, 18th - 21st centuries, <https://ristat.org/>, Version I (2021): Datatype 7.01 - Land, benchmark-year 2002.

or, for the documentation to a data-set:

Kessler, Gijs and Andrei Markevich, Electronic Repository of Russian Historical Statistics, 18th - 21st centuries, <https://ristat.org/>, Version I (2021): [Title] [page numbers] [filename: XXXXX]

Example:

Kessler, Gijs and Andrei Markevich, Electronic Repository of Russian Historical Statistics, 18th - 21st centuries, <https://ristat.org/>, Version I (2021): Modern Classification: Principles and Documentation - Land, p. 1 [filename: ERRHS_7_00_Modern_Classification_EN.pdf]