Bulletin USAMV-CN, 63/2007

NATIONAL TRENDS ON AGRICULTURAL CROPS PRODUCTION: CLUSTER ANALYSIS

Jäntschi L[1], Sorana D. Bolboacă[2], Carmen E. Stoenoiu[1]

[1]Technical University of Cluj-Napoca, 103-105 Muncii Bvd., 400641 Cluj-Napoca, Romania, lori@academicdirect.org, carmen@j.academicdirect.ro
[2]"Iuliu Hatieganu" University of Medicine and Pharmacy, 6 Louis Pasteur, 400349 Cluj-Napoca, Romania, sorana@j.academicdirect.ro

Key words: Resources Management; Crops Production; National Trends; Cluster Analysis

Abstract
Staring from descriptive data on crop production and cultivated area at national level during on fifteen years, the aim of this study is to reveal the trends on crops cultivation. The cluster analysis reveals linkages between crops classes as well as between different crops, which can be partly assigned to crops rotation. Time series analysis reveals dramatically reducing of production of some crops, such as flax, hemp, and sugar beet, and increasing of production, such at sunflower, and increasing of productivity, such at potatoes and field vegetables.

INTRODUCTION

The Romanian agriculture used to have a major contribution to Romanian economy but unfortunately the return of collectivized farmland to its cultivators, one of the first initiatives of the post-December 1989 revolution government, resulted in a short-term decrease in agricultural production. Some four million small parcels representing 80% of the arable surface were returned to original owners or their heirs (Dawidson, 2005). Many of the recipients were elderly or city dwellers, and the slow progress of granting formal land titles was an obstacle to leasing or selling land to active farmers. The contribution of agriculture to the Gross National Product (GNP) in Romania was around 13.7% in 1989, increasing to 18.6% in 1995 and decreasing to 12.9% in 1999 and 11.4% in 2000 respectively (MARD, 2001).

The major crops produced in Romania are: wheat and rye, barley and two-row barley, corn, soya, sunflowers (Turtoi et al., 2004), oats, rice, hay, potatoes, soybeans, sugar beets, feed roots, corn silage, and tobacco (Bachman, 1989). Today, the crops production must be aligned to the policy of European Union (Chantreuil et al., 2005; Hertel et al., 1997) and to the climate changes (Cuculeanu et al., 1999).

Based on descriptive information about cultivated area and total production by year and by crop the aim of the study is to identify national trends in crops production, as well as linkages between productions of different crops.

MATERIAL AND METHOD

Material

Nine main crops cultivated in Romania on a period of fifteen years (from 1990 to 2004) have been investigated. Data were taken from administrative sources: National Administrative of Land Improvement, for Agricultural Irrigated Area and Ministry of Agriculture, Forests

and Rural Development (Romanian Statistical Yearbook, 2005). The investigated main crops were: ***Cereals for grains:*** wheat and rye, barley and two-row barley, oats, maize, sorghum, rice; ***Leguminous crops for beans:*** peas beans, bean; ***Industrial crops:*** *Fiber crops:* flax for fiber, hemp for fiber; *Oilseed crops:* sunflower, rape, soya beans, flax for oil, castor plant; *Other industrial crops:* sugar beet, tobacco, medicinal and aromatic plants; ***Potatoes:*** autumn potatoes; ***Vegetables:*** tomatoes, dry onion, dry garlic, cabbage, green peppers, edible roots; ***Water melons and melons***; ***Fodder crops:*** old and new perennials, lucerne, clover, *annuals for hay and green fodder*; *Plants used for silage:* maize for silage, *fodder roots*; ***Total Fruits:*** plums, apples, pears, peaches, cherries and sour cherries, apricots and engrafted apricots, nuts, strawberries, and other fruits.

Method

A time series is a sequence of data points, measured typically at successive times, spaced at (often uniform) time intervals. Note that our data fits to this definition. Time series analysis comprises methods that attempt to understand such time series, often either to understand the underlying theory of the data points (what generated them?), or to make forecasts (predictions). Time series coming from earth and life sciences study include trend, cyclicity, and periodicity (Jäntschi, 1995).

Clustering is the classification of objects (our objects are crops) into different groups, or more precisely, the partitioning of a data set into subsets (named clusters), so that the data in each subset (ideally) share some common trait - often proximity according to some defined distance measure. Data clustering is a common technique for statistical data analysis, which is used in many fields, including data mining and bioinformatics (Bolboacă and Jäntschi, 2007).

We used both time series analysis and cluster analysis for our purpose of trend and linkages identification on national crop production. The subjects of analysis were cultivated area, and crop production for main crops.

RESULTS AND DISCUSSIONS

The Romanian cultivated area, by main crops, expresses as the coefficient obtained by division between the cultivated area in one year and cultivated area in the previous year is presented in Table 1.

Table 1. Cultivated area by main crops: comparison with previous year

Crop \ years	1991/ 1990	1992/ 1991	1993/ 1992	1994/ 1993	1995/ 1994	1996/ 1995	1997/ 1996	1998/ 1997	1999/ 1998	2000/ 1999	2001/ 2000	2002/ 2001	2003/ 2002	2004/ 2003
Cereals for grains	1.06	0.95	1.11	1.03	0.98	0.91	1.08	0.94	0.91	1.05	1.11	0.96	0.92	1.13
Leguminous crops for beans	0.63	0.85	0.96	1.01	0.94	1.07	0.79	0.84	1.03	0.90	0.86	1.27	1.03	0.82
Industrial crops	1.00	1.16	0.80	0.98	1.18	1.22	0.87	1.27	1.01	0.85	0.88	1.14	1.27	0.84
Potatoes	0.81	0.93	1.14	1.00	0.98	1.05	0.99	1.02	1.05	1.03	0.98	1.02	1.00	0.94
Vegetables	0.90	1.14	0.99	0.93	1.05	1.02	0.96	1.07	1.04	1.00	0.98	1.03	1.02	1.27
Water melons and melons	1.42	0.97	0.95	0.93	1.11	1.07	0.86	1.05	1.11	0.93	0.84	1.12	0.97	0.89
Fodder crops	0.79	0.93	0.91	0.95	0.97	1.02	0.91	1.01	1.03	0.94	0.93	1.18	1.06	0.44
Plants used for silage	0.42	1.16	0.62	0.79	0.86	1.14	0.61	1.07	0.67	0.88	0.71	1.34	0.83	0.79
Total	0.98	0.97	1.03	1.01	1.00	0.96	1.02	0.99	0.95	1.00	1.05	1.01	0.99	0.96

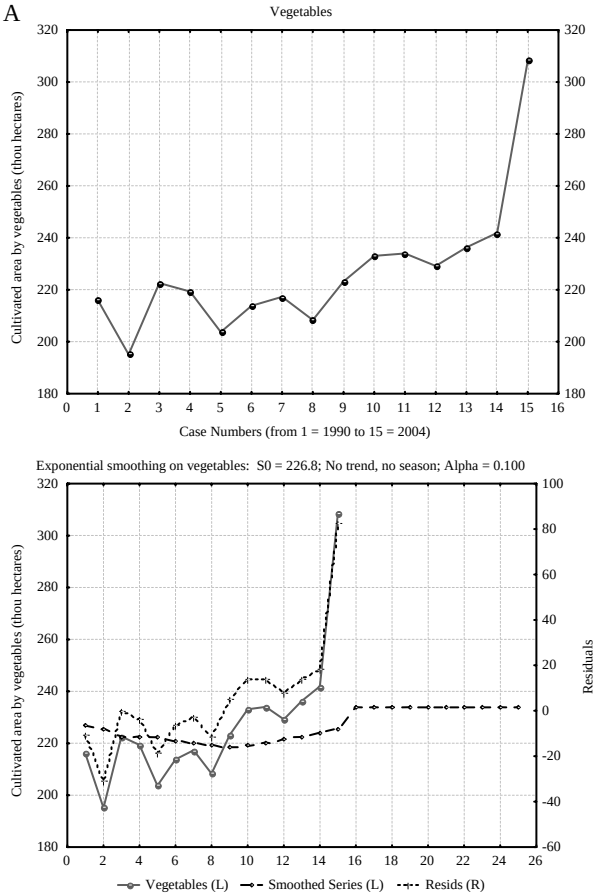
The date regarding the total crops production by class was summarized in Table 2.

Table 2. Characteristics of total crops production

coef [year_x/year_x-1] / Crops class	1991/ 1990	1992/ 1991	1993/ 1992	1994/ 1993	1995/ 1994	1996/ 1995	1997/ 1996	1998/ 1997	1999/ 1998	2000/ 1999	2001/ 2000	2002/ 2001	2003/ 2002	2004/ 2003	Min	Max	Mean
Cereal grains	1.12	0.64	1.26	1.17	1.09	0.71	1.56	0.70	1.10	0.61	1.80	0.76	0.90	1.88	10477	24403	16813
Leguminous crops for beans	0.71	0.94	1.14	0.89	1.27	0.79	1.02	0.92	1.06	0.48	1.66	0.90	1.10	1.85	37	112	77
Oilseed crops	1.11	1.12	0.89	1.06	1.21	1.15	0.82	1.32	1.22	0.54	1.16	1.19	1.47	1.13	739	1995	1147
Crops for other industrial purposes	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	0.62	1.31	1.18	0.71	1.40	19	30	25
Potatoes	0.59	1.39	1.43	0.79	1.02	1.19	0.89	1.04	1.19	0.88	1.15	1.02	0.97	1.07	1873	4230	3409
Vegetables – total	0.94	1.19	1.09	0.89	1.12	0.95	0.89	1.16	1.08	0.83	1.14	1.00	1.17	1.42	2214	4774	2863
Water melons and melons	1.94	0.84	0.97	1.02	1.05	1.09	0.90	1.10	1.24	0.62	1.04	1.18	1.17	1.00	382	853	648
Total fruit	0.80	1.00	1.87	0.45	0.94	1.78	0.87	0.73	0.90	1.39	1.04	0.70	2.19	0.84	917	2182	1355

Coef[year_x/(year_x-1)] = (total production on year_x)/(total production on (year_x-1)];
Min = minim production (thou tones); Max = maximum production (thou tones); Mean = average production for fifteen years

A



B

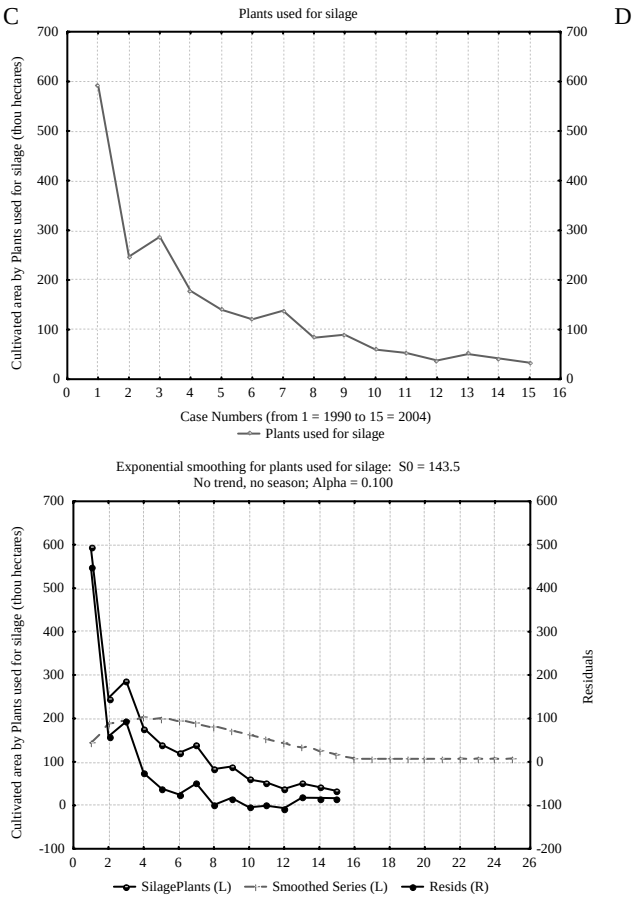


Figure 1. Time series analysis on cultivated area by vegetables and plants used for silage: (A) Spectral Fourier plot for vegetables (mean error 4.73); (B) Exponential smoothing chart for vegetables; (C) Spectral Fourier plot for plants used for silage; (D) Exponential smoothing chart for plants used for silage

Cultivated Area

Analyzing the coefficients on each main crop from Table 1 the following are true:

- The best years (in terms of cultivated area with the main crops): 1996 and 2002 (in seven out of eight cases the surface of the cultivated area by main crops was increased comparing with 1995, and 2001 respectively);
- The worst years in terms of cultivated area with the main crops: 1997, and 2001 (just in one case the surface cultivated with cereals for grains were increased comparing with 1996, and 2000, respectively);
- Top three most cultivated area by class of crops: vegetables (starting with 1995 the surface cultivated with vegetables increased comparing with previous year, excepting the years 1997, and 2001), industrial crops (the surface cultivated increased in eight years comparing with previous years); cereals for grains (the surface cultivated increased in seven years comparing with previous years);
- Less cultivated area by class of crops: plants used for silage (just in four years the surface was increased comparing with the previous years).

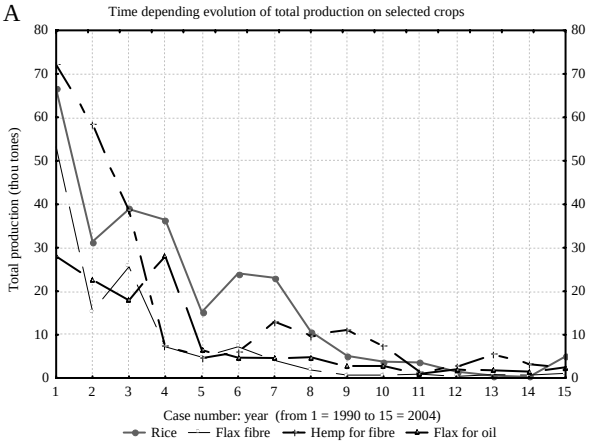
Analyzing the overall tendency, it can be observed that, for example the worst years in terms of cultivated areas by the main crops registered an increasing comparing with previous years, this being done in the privilege of one crop and underprivileged all others main crop.

Thus, in 1997 reported to 1996, the total surface cultivated with wheat and rye increased with 626.7 thou hectares, and the total surface cultivated with barley and two-row barley

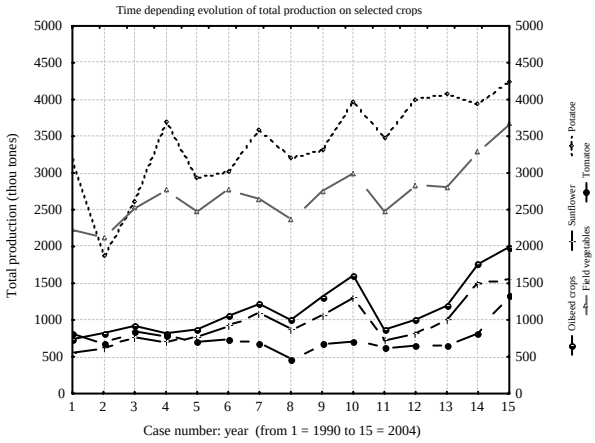
increased with 111.1 thou hectares, while surface cultivated by all others main crops decreased.

A time series analysis of one of the top class of cultivated crops by cultivated area and opposite of the less cultivated class of crops was performed and the results are presented in Figure 1. Note that there was not identified any seasonality on any of the cultivated surface on the main investigated crops.

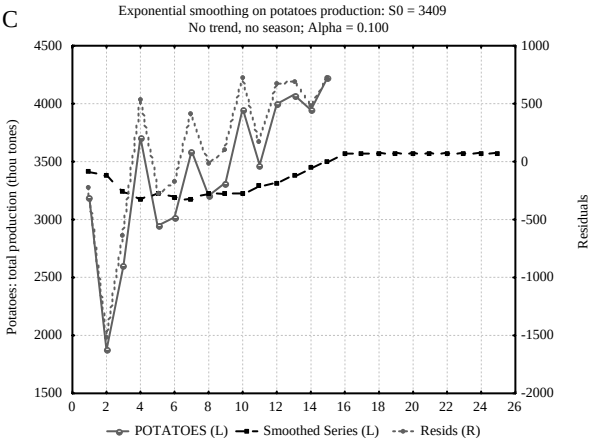
A



B



C



D

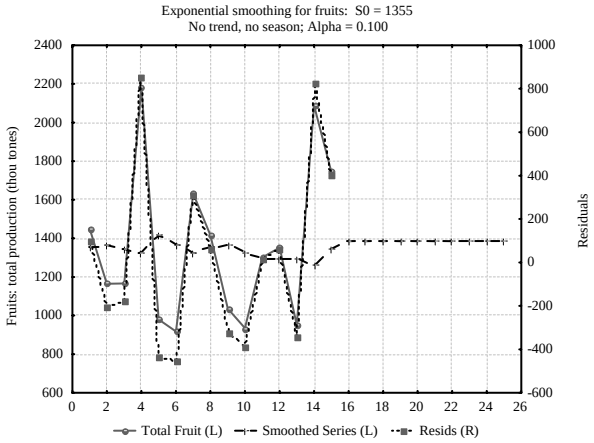


Figure 2. Time series analysis on total production by selected crops: A. Main crops with decreasing production trend; B. Main crops with increasing production trend; C. Exponential smoothing chart on potatoes production; D. Exponential smoothing chart on fruits production

The goodness-of-fit of the forecast models (Figure 2 B and D) are sustained by the values of mean absolute error, mean percent error and mean absolute percent error values. The forecast models for cultivated area by vegetables and silage plants are more performing at the beginning and at the end of the investigated years on cultivated areas by vegetables as well as on cultivated areas by silage plats.

Total Crops Production

By analyzing the types of crops in classes the following were identified:

- In almost twenty-three percent of cases, the highest production was obtained in 2004: wheat and rye, maize, sorghum, peas beans, oilseed crops, sunflower, soybean, autumn potatoes, field vegetables, tomatoes, and apples. It can be said that 2004 was a good year on main crop production for Romania.
- In ten cases a reduction of the total production, associated with a maximum production in first year analyzed (1990): rice, bean, flax fiber, hemp for fiber, annuals for hay and green fodder, plants used for silage, maize used for silage, fodder roots, peaches, and apricots and engrafted apricots.

A number of main crops were identified as having an ascendant (the total production was increased from 1990 to 2004) or descending trend (the total production was decreased starting with 1990 to 2004) (see figure 2 A and B).

Analyzing the data presented in table 2 the followings are true:

- The top three highest total production were for the registered for: oilseed crops (in eleven cases the total production increased reported to previous investigated year), water melons and melons (in ten cases the total production increased reported to previous investigated year), and potatoes (in nine cases the total production increased reported to previous investigated year). Note that the highest production is not related to the largest cultivated area (see previous subsection of Results and Discussion section);
- The less total production by the main crops: total fruits (just in six out of the investigated years the total production was higher comparing with previous investigated year).

A seasonal and non-seasonal analysis has been performed in order to identify any seasonality on vegetables and fruits productions on investigated years (see Figure 2 C and D). There was not identified any seasonality on any of the main investigated crops regarding the total production.

The goodness-of-fit of the forecast (Figure 2 C and D) models are sustained by the values of mean absolute error, mean percent error and mean absolute percent error values. Note that the forecast models are for total fruits production is more closed on the total production data comparing with potatoes; the residuals varied in a larger scale for potatoes comparing with total fruits.

Clusterization Analysis on Cultivated Areas and Total Productions

Hierarchical cluster analysis technique has been applied in order to identify similarities on crop classes (Table 3, left), and on main crops (Table 3, right). Seven cases out of nine were valid for analysis on crops classes, having enough data for analysis on both cultivated area and production; twenty-nine out of thirty-six cases were valid on main crops analysis for cultivated area and cultivated area together with total production data analysis.

Analyzing the dendrograms presented in Table 3, the followings are true:

- The cereals for grains (entries #, §, and ¥) does not join with any other crop into a cluster until the highest distance on all of the included parameters (cultivated area - cluster W, total production - cluster Y, and both of them - cluster Z). This fact shows that cereals for grain are the main class of interest, having significantly larger cultivated area, significantly bigger production.
- For all crops classes excepting the cereals for grains, cultivated area and total production cauterizes together (clusters V, $, and X), all at lowest distance (1), meaning that cultivated area and total production characteristics are near one to each other these.

Cultivated area		Both parameters		Total production		▼ Crop ▼		Total production						Both parameters						Cultivated area			
#	W	§	Z	¥	Y	Cereals for grains	wheat & rye				F	G	H					N	O	P			T	U
							barley	C	D	E			H	K	L	M			P	R	S		U
							oats	A		E			H	I		M			P	Q	S		U
							maize					G	H					O	P			T	U
							sorghum	A		E			H	I		M			P	Q	S		U
							rice	A		E			H	I		M			P	Q	S		U
V	W	$	Z	X	Y	Leguminous crops for beans	peas beans	A		E			H	I		M			P	Q	S		U
							bean	A		E			H	I		M			P	Q	S		U
V	W	-	-	X	Y	Industrial crops	flax for fiber	A		E			H	I		M			P	Q	S		U
							hemp for fiber	A		E			H	I		M			P	Q	S		U
							sunflower	A		E			H	I		M			P	R	S		U
							rape	A		E			H	I		M			P	Q	S		U
							soya beans	A		E			H	I		M			P	Q	S		U
							flax for oil	A		E			H	I		M			P	Q	S		U

							sugar beet	C	D	E			H	K	L	M			P	Q	S		U								
							tobacco	A		E			H	I		M			P	Q	S		U								
							medic.& arom.	A		E			H	I		M			P	Q	S		U								
V	W	$	Z	X	Y	Potatoes	autumn ones	B	D	E			H	J	L	M			P	Q	S		U								
V	W		Z	X	Y	Vegetables	tomatoes	A		E			H	I		M			P	Q	S		U								
							dry onion	A		E			H	I		M			P	Q	S		U								
							dry garlic	A		E			H	I		M			P	Q	S		U								
							cabbage	A		E			H	I		M			P	Q	S		U								
							green peppers	A		E			H	I		M			P	Q	S		U								
							edible roots	A		E			H	I		M			P	Q	S		U								
V	W	$	Z	X	Y	Water melons and melons		A		E			H	I		M			P	Q	S		U								
V	W	-	-	X	Y	Fodder crops	lucerne				F	G	H				N	O	P	Q	S		U								
							clover	B	D	E			H	J	L	M			P	Q	S		U								
							maize for silage	C	D	E			H	K	L	M			P	Q	S		U								
							plums	A		E			H	-	-	-	-	-	-	-	-	-	-								
V	W	-	-	X	Y	Total Fruits	apples	A		E			H	-	-	-	-	-	-	-	-	-	-								
							pears	A		E			H	-	-	-	-	-	-	-	-	-	-								
							cherries	A		E			H	-	-	-	-	-	-	-	-	-	-								
							apricots	A		E			H	-	-	-	-	-	-	-	-	-	-								
							nuts	A		E			H	-	-	-	-	-	-	-	-	-	-								
							strawberries	A		E			H	I		M			P	Q	S		U								
							other fruits	A		E			H	-	-	-	-	-	-	-	-	-	-								
1	25	1	25	1	25	◄ Distance ►		1	2	3	4	8	25	1	2	3	5	9	25	1	2	4	25								

Table 3. Dendrogram on Crops vs. Rescaled distance cluster combine using Average linkage between groups

- Total production of main crops are similarly (cluster A) with some exceptions. B cluster strong relates (at distance 1) in terms of production clover with autumn potatoes, one of the possible explanations being that the rotation of these two crops being recommended (Wikipedia, 2007). C cluster strong relates maize for silage with sugar beet, and here we can found in the literature a paper which recommend the rotation between these two (Maas and Lamers, 1988). Remaining crops clusterizes as follows: wheat and lucerne - cluster F, at distance 4 (which is recommended by many authors as rotating culture for cereals - Ward et al., 2003), and both together with maize (cluster G, at distance 8), which is recommended rotating culture for wheat.
- Similar associations reveals Table 3 on cultivated area parameter combined with total production parameter. On Cultivated area parameter, lucerne leaves the cluster of wheat & rye and maize (cluster T, at distance 4).

CONCLUDING REMARKS

Cluster analysis as well as time series analysis reveled that cereals for grains are an apart class of crops in terms of cultivated area and production at national level, having not only the most important part at national level, but also being totally different from all others.

Rotation of cultures has also an important influence on crop production and cultivated areas at national level, as were underlined for wheat, lucerne and maize.

Some cultures have been almost abandoned, such as flax, hemp, and sugar beet. Other cultures such as potatoes and field vegetables increases in total production, even if it decreases in cultivated area. Sunflower is one of few cases of increasing of both cultivated area and productivity, over 95% being produced by private majority ownerships in 2004.

ACKNOWLEDGMENTS

This research was partly supported by UEFISCSU Romania through project ET46/2006 and CNCSIS Romania through project AT93/2007.

REFERENCES

1. ***, Agriculture and Sylviculture (Chapter 14th) [online document]. Romanian Statistical Yearbook, 2005, Available from: URL: http://www.insse.ro/cms/files/pdf/ro/cap14.pdf.
2. ***, Ministry of Agriculture and Rural Development - Document of the Romanian position - Chapter 7 - Agriculture, 2001, located at: http://mapam.ro/pages/ue/cap7-agricultura.pdf, accessed August 20, 2007.
3. Dawidson KEK, 2005, Redistribution of land in post-communist Romania, Eurasian Geography and Economics, 46(8), 618-632.
4. Bachman RD, 1991, Area Handbook Series: Romania: A Country Study, Library of Congress, Washington DC, 1-356.
5. Sorana D Bolboacă, L Jäntschi, 2007, Data Mining on Structure-Activity/Property Relationships Models, World Applied Sciences Journal, 2(4), pp. 323-332.
6. Chantreuil F, F Levert, E Erjavec, T Donnellan, K. Hanrahan, 2005, CAP reform: Prospects for crop markets in the Enlarged EU, EuroChoices 4(1), 26-31.
7. Jäntschi L, 1995, Time series. Prediction (in Romanian) - BSc Thesis in Informatics (advisor: Frențiu M.), Babeş-Bolyai University, available online.
8. Turtoi C, G Cornélissen, D Wilson, F Halberg, 2004, About 21-year Hale cyclic signature in agriculture complement those in biomedicine, Biomedicine and Pharmacotherapy, 58(suppl. 1), S12-S14.
9. Cuculeanu V, A Marica, C Simota, 1999, Climate change impact on agricultural crops and adaptation options in Romania, Climate Research, 12(2-3 spec. iss. 6), 153-160.
10. Maas PWTH, JG Lamers, 1988, Management of the yellow beet cyst nematode with crop rotation, soil fumigation and granular nematicides, European Journal of Plant Pathology, 94(3), 113-122.
11. Hertel TW, M Brockmeier, PV Swaminathan, 1997, Sectoral and economy-wide analysis of integrating Central and Eastern European countries into the EU: Implications of alternative strategies, European Review of Agricultural Economics, 24(3-4), 359-386.
12. Wikipedia, 2007, Examples of crop rotation / Agriculture - from article “Crop rotation” (http://en.wikipedia.org/wiki/Crop_rotation), Created on August 12, 2007, cited on August 30, 2007.
13. Ward P, P Dolling, F Dunin, 2003, The impact of a lucerne phase in a crop rotation on groundwater recharge in south-west Australia, Invited presentation, 11th Australian Agronomy Conference: Solutions for a new environment, Geelong, February 2003, online (http://www.regional.org.au/au/asa/2003/i/5/ward.htm).