Foresight of Macro Environment in Agribusiness: Dynamic Relationships of Food Consumption and Agricultural Production (Analysis of the Relations between Agricultural Production and Domestic Consumption)

Roman B. SHESTAKOV^{1*}, Nikolay A. YAKOVLEV¹, Galina P. ZVEREVA¹,
and Anastasiya S. VOLCHENKOVA¹

¹ Orel state agrarian University named after N.V. Parahin, Oryol, Russian Federation

* Correspondence: satynter@mail.ru

Abstract: The article is devoted to an important economic element - agribusiness-foresight - the analysis of general economic conditions, and specifically the search for links in the production of agricultural products in general and domestic consumption of basic food products. As the main research methods for the current research there were used LASSO Regression method with a regularizer and Gradient Boosting Machine algorithm. The time series from 1991 to 2018, transformed into the form of annual percentage changes, were used as the initial data. At the beginning, a preliminary selection of features was made, and further two corresponding models were formed on the basis of the selected hyperparameters. As an additional target variable, a series of production was created with a time lag of one period in relation to the rest of the series. The degrees of the linear and nonlinear relationships for various types of food are determined, as well as differences in aggregates of crop production and animal husbandry are revealed. General recommendations are given within the framework of the policy for the development of the institutional environment of agribusiness and its foresight systems.

Keywords: consumption of agribusiness, production index, food, agriculture, relationships, foresight.

1. Introduction

Modern agribusiness management functions go beyond simplified planning and forecasts, approaching full-fledged foresight systems [1, p. 48-54] for decision-making in a wide range of business activities. Further digitalization, the spread of the Internet of things (including means of production) help to reveal the internal potential of the enterprise for effective activities [2]. At the same time, taking into account economic difficulties and uncertainty, especially the functioning of the domestic agro-industrial complex, it is equally important to have a long-term view of economic development in retrospect (ex post) and in the future (ex ante). There is a necessity of a fundamental analysis of the environment where the subjects operate according to different parameters.

Food consumption is important in the point of view of two general business positions. The first position is a reflection of the level of food security, which depends not only on covering domestic needs with food, but also on their availability to the general public. The second one is related to the development of domestic consumption as such, and the growth of domestic consumption is the cornerstone of economic growth in the country. The leading role here remains with the agricultural sector, since food, in fact, is the main result of its functioning.

Thus, the urgent task of the study is to identify the interactions and their level between food consumption and agricultural production within the national economic system.

2. Methods and Materials

It is supposed to use data of basic foodstuff consumption in the Russian Federation in physical terms as the initial features (independent variables), and to use the production index of agricultural productivity, also in the whole country (Table 1.2), as the target attribute. To facilitate the manipulation of these data later in the text, we assign them a simplified letter code.

Table 1. Consumption of basic food products in the Russian Federation (per capita per year; kilograms) [3]

Products	Code	1990	1995	2000	2005	2010	2015	2016	2017	2018
Potato	A	106	124	109	109	95	91	90	90	89
Vegetables and gourds	B	89	76	79	87	98	102	102	104	107
Fruits and berries	C	35	29	32	46	57	60	60	59	61
Meat and meat products in terms of meat	D	75	55	45	55	69	73	74	75	75
Milk and dairy products in terms of milk	E	387	254	215	234	245	233	231	230	229
Eggs and egg products, pieces	F	297	216	229	250	270	268	273	279	280
Sugar	G	47	32	35	38	39	39	39	39	39
Vegetable oil	H	10,2	7,5	9,9	12,1	13,4	13,6	13,7	13,9	14,0
Bread Products (Bread and pasta in terms of flour, flour, groats and beans)	K	120	122	117	121	120	118	117	117	116

Table 2 presents the values of the indices of agricultural production in the Russian Federation.

Table 2. Indices of agricultural production by categories of farms in the Russian Federation (in comparable prices; percent of the previous year) [3]

Year	Code	1990	1995	2000	2005	2010	2015	2016	2017	2018
Value	Y	96,4	92,0	106,2	101,6	87,9	102,1	104,8	102,9	99,8

Thus, there were obtained an initial "date frame" consisting of 29 observations (years), 9 attributes or independent variables (types of main products) and a target variable in the form of an agricultural production index. To level out the occurrence of false correlation and regression of time series, we transform the signs into the series of percentage changes over the years (growth). The target variable, in fact, is already a percentage change (growth); in order to find the gain it was reduced by 100. Then there were checked the feasibility of using the time lag from Y at the input (model with distributed lag) and at the output, as an additional target variable. A series shift was by 1 period so that future values Y_{t+1} (Y lag) correspond to independent variables.

Table 3. Indicators transformed into a series of percentage changes

Year	A	B	C	D	E	F	G	H	K	Y	Y lag
1991	0,07	-0,03	0,00	-0,08	-0,10	-0,03	-0,19	-0,24	0,01	-4,50	-9,40
1992	0,04	-0,10	-0,09	-0,13	-0,19	-0,09	-0,21	-0,14	0,03	-9,40	-4,40
1993	0,08	-0,08	-0,09	-0,02	0,04	-0,05	0,03	0,04	-0,02	-4,40	-12,00
1994	-0,03	-0,04	-0,03	-0,03	-0,04	-0,05	0,00	-0,04	0,01	-12,00	-8,00
1995	0,01	0,12	0,04	-0,04	-0,10	-0,09	0,03	0,12	-0,02	-8,00	-5,10
1996	-0,01	-0,03	0,07	-0,07	-0,08	-0,04	0,03	0,07	-0,03	-5,10	0,90
1997	-0,01	0,03	0,06	-0,02	-0,01	0,01	0,00	0,05	0,00	0,90	-14,10
1998	-0,07	-0,03	-0,09	-0,04	-0,04	0,03	0,00	0,05	-0,01	-14,10	3,80
1999	-0,05	0,07	-0,10	-0,06	-0,03	0,02	0,03	0,06	0,02	3,80	6,20
2000	0,01	0,00	0,19	0,00	0,00	0,04	0,03	0,06	-0,02	6,20	6,90
2001	0,00	0,03	0,09	0,04	0,02	0,03	0,00	0,05	0,03	6,90	0,90
2002	-0,03	-0,01	0,11	0,06	0,04	0,04	0,03	0,01	0,01	0,90	-0,10
2003	0,03	0,05	0,00	0,04	0,02	0,01	0,00	0,05	-0,01	-0,10	2,40
2004	-0,01	0,01	0,10	0,04	0,00	-0,01	0,03	0,05	-0,01	2,40	1,60
2005	0,01	0,02	0,07	0,02	0,01	0,03	0,03	0,04	0,02	1,60	3,00

Year	A	B	C	D	E	F	G	H	K	Y	Y lag
2006	0,01	0,03	0,04	0,05	0,01	0,02	0,03	0,03	0,00	3,00	3,00
2007	-0,01	0,02	0,04	0,05	0,01	0,00	0,00	0,02	-0,01	3,00	10,60
2008	-0,03	0,08	0,06	0,08	0,00	-0,01	0,00	0,00	-0,01	10,60	0,70
2009	-0,01	0,02	0,04	0,00	0,01	0,03	-0,05	0,02	-0,01	0,70	-12,10
2010	-0,10	-0,03	0,04	0,05	0,01	0,04	0,05	0,03	0,02	-12,10	22,30
2011	0,01	0,04	0,04	0,03	-0,01	0,01	0,03	0,01	-0,01	22,30	-5,60
2012	0,01	0,01	0,02	0,04	0,01	0,01	0,00	0,01	0,00	-5,60	5,10
2013	-0,03	-0,01	0,05	0,01	0,00	-0,02	0,00	0,00	-0,01	5,10	4,10
2014	-0,01	0,00	0,00	-0,01	-0,02	-0,01	0,00	0,01	0,00	4,10	2,10
2015	-0,02	0,00	-0,05	-0,01	-0,03	0,00	-0,03	-0,01	0,00	2,10	4,80
2016	-0,01	0,00	0,00	0,01	-0,01	0,02	0,00	0,01	-0,01	4,80	2,90
2017	0,00	0,02	-0,02	0,01	0,00	0,02	0,00	0,01	0,00	2,90	-0,20
2018	-0,01	0,03	0,03	0,00	0,00	0,00	0,00	0,01	-0,01	-0,20	-

Selection of variables for the model was performed in the following way. The first step is to remove the variables with the minimum standard deviation. In our case, it is K - bread products. Next, we construct the correlation matrix by the Pearson method (Table 4).

Table 4. The Correlation matrix of the used features

	A	B	C	D	E	F	G	H	Y
B	-0,2								
C	-0,1	0,3							
D	-0,2	0,4	0,5						
E	-0,2	0,3	0,4	0,8					
F	-0,4	0,2	0,3	0,6	0,7				
G	-0,5	0,4	0,3	0,6	0,7	0,4			
H	-0,4	0,5	0,3	0,4	0,5	0,3	0,9		
Y	0,1	0,5	0,4	0,5	0,4	0,3	0,2	0,1	
Y_{t+1}	-0,6	0,1	0,2	0,4	0,2	0,4	0,4	0,2	0,0

It is acceptable to select features with a correlation of at least 0.5 for the target variable. However, some variables with a lower correlation in the analysis of paired scatter plots show the presence of nonlinear relationships. We also note that significant autocorrelation in the Y series is not observed. This additionally confirms the ARIMA

analysis, identifying it as an integrated series with a moving average of the second order (0, 1, 2) [4]. Also, the correlation with the independent variables in the "lagged" series Y_{t+1} is less pronounced than with the original Y . At this stage, only the characteristic variables A , H , Y_{t+1} are removed from the proposed model. However, Y_{t+1} is left as an additional target variable. In addition, it is noted that the remaining signs are observed multicollinearity (cross-correlation > 0.5). There are chosen two models that are resistant to the above conditions for comparison: LASSO Regression and the Gradient Boosting Machine algorithm. LASSO is resistant to multicollinearity and can be used to select meaningful data. GBM is based on sequential training of decision trees and will help to cover non-linear relationships and also rank attributes by importance.

3. Results and Discussion

The current models will not be used for forecasting, so there was no additional separation of the "hold-out" test sample, in favor of cross-validation on the entire data set (5 splits). The main task remains to study the effect of input consumption variables on the dynamics of agricultural production. Model fitting was performed using specialized Sklearn and Xgboost modules in the Python software environment [5, 6].

Based on the calculation results, the following parameters were obtained (Table 5).

Table 5. Comparison of the obtained model parameters

Features	LASSO Regression		GBM	
	Y	Y_{t+1}	Y	Y_{t+1}
Hyper Parameters				
1	alpha: 0,15	alpha: 0,05	n_estimators: 10	n_estimators: 11
2	Intercept = False	Intercept = False	learning_rate: 0,05	learning_rate: 0,05
3	-	-	colsample_bytree: 0,7	colsample_bytree: 0,2
4	-	-	gamma: 2	gamma: 3
5	-	-	max_depth: 5	max_depth: 5
Parameters	Coefficients		Values of features	
E	0	0	0,32	0,20
B	0	0	0,27	0,23
G	0	1,51	0,21	0,25
F	0	0	0,09	0,16
C	9,42	0	0,07	0
D	0,08	0	0,04	0,15

As for a linear model Y with a regularizer, the only coefficients greater above zero are C and D , and C is most significant. However, the quality of the model taking into account R^2 is very low, and explains only 6% of Y variability. The approximation of the Y_{t+1} model is even worse and is not worth further consideration.

More interesting results are shown by GBM. Value of features is opposite to the previous model, perhaps not least due to the capture of non-linear patterns in the data. The most significant are E , B , G , just below are F , C , D . The situation is similar in the model taking into account the time lag. More specifically, within the framework of coinciding periods, the following products have the greatest connection with the dynamics of agricultural production: milk and dairy products, sugar, vegetables and gourds. The connection with the next year of production in these groups remains at a similar level, and it increases for meat and meat products, eggs and egg products, but for fruits and berries it practically disappears.

In other words, the consumption of livestock products shows an inertial reaction, which is a consequence of the livestock industry characteristics [7, p. 5-19]. Thus, in the current economic situation it can be summarized that domestic consumption is not a strong driver of agricultural production for most products. The dynamics of consumption in most food time series in itself did not possess noticeable variability and growth trends, and with a decrease in the macroeconomic consumer background, it is likely to decline.

At the institutional level, the agro-industrial complex management strategy as part of the implementation of state development programs requires adjustment [8, p. 82-87], including the foregoing view. Highly processed export support should play an important role here. A more intensive implementation of digital elements in the supply chain, integrated foresight and the integration of science, technology and innovation [9] will help increase the efficiency of agribusiness. The flip side of the process will be the explosive growth of data requiring processing and interpretation [10]. Improving the systems of long-term planning and decision-making based on the analysis of big data and risks is impossible without the involvement of the academic environment and the expert community [11, p. 265-266.], as well as strengthening of local public-private partnerships. This is especially true in the context of technological sanctions.

The results of this work also outline further research tasks that may involve a deeper study of domestic consumption, the influence of price factors, and export and import analysis. It is also useful to assess the role of production of the non-food component in agriculture, to comprehensively study the complex system of mutual influence of related areas of the agricultural sector at the regional and federal levels.

4. Conclusion

The article provides a regression analysis to determine the relationship between the consumption of in-kind basic food products and the dynamics of agricultural production from 1991 to 2018. All time series are transformed to the form of annual percentage changes; preliminary selection of attributes is made. As a response, the production line was additionally compared with the time lag. A linear model of LASSO Regularizer and GBM were used. It is determined that in the majority of their relationships are not linear in nature. The effect varies by food group. For example, livestock products have a greater inertia in time. It can be concluded that the impact of domestic consumption on the

dynamics of agricultural production is insufficient and requires additional attention in the framework of the development policy of the institutional environment of agribusiness as a whole and its foresight systems.

References

- (1) Shestakov, R. B.; Yakovlev, N. A. Some of the Forecast Function Aspects in an Agricultural Production Foresight-Process (in Russian). *International Scientific Journal* **2019**, 3, 48–54. <https://doi.org/10.34286/1995-4638-2019-66-3-48-54>.
- (2) Tzafestas, S. G. The Internet of Things: A Conceptual Guided Tour. *European Journal of Advances in Engineering and Technology* **2018**, 5 (10), 745–767.
- (3) Official site of RosStat <https://eng.gks.ru/> (accessed Feb 22, 2020).
- (4) Forecasting functions for time series and linear models <https://www.rdocumentation.org/packages/forecast/versions/8.9> (accessed Feb 22, 2020).
- (5) Scikit-learn: machine learning in Python <https://scikit-learn.org/stable/> (accessed Feb 22, 2020).
- (6) XGBoost Documentation <https://xgboost.readthedocs.io/en/latest/> (accessed Feb 22, 2020).
- (7) Buyarov, V. S.; Gudymenko, V. I.; Buyarov, A. V. *Economics and Reserves of Poultry Meat: A Monograph* (in Russian); Publishing House of the Federal State Budget Educational Establishment of Higher Education of Oryol State Agricultural University: Orel, 2016.
- (8) Solodovnik, A. I.; Lovchikova, E. I.; Fedotenkova, O. A.; Hashir, A. A. Strategic Terms of Management by Agro-Industrial Complex within the Framework of Implementation of the State Development Programs (in Russian). *Economy, Labor, Management in Agriculture* **2019**, No. 2, 79–87. <https://doi.org/10.33938/192-79>.
- (9) Bharadi, V. A.; Abhyankar, P. P. Analysis and Prediction in Agricultural Data Using Data Mining Techniques. *International Journal of Research in Science & Engineering* **2017**, Special Issue 7-ICEMTE, 386–393.
- (10) Channe, H.; Kothari, S.; Kadam, D. Multidisciplinary Model for Smart Agriculture Using Internet-of-Things (IoT), Sensors, Cloud-Computing, Mobile-Computing & Big-Data Analysis. IJCTA May-June 2015. *International Journal of Computer Technology and Applications* **2018**, 6 (3), 374–382.
- (11) Shestakov, R. B.; Yakovlev, N. A.; Tarakin, A. V. On the Issue of Monitoring and Forcing of the Regional Agro-Industrial Complex in the Era of Digital Transformation. *Innovations and Investments* **2018**, 12, 265–266.

Information about Authors

Roman Borisovich SHESTAKOV: Ph.D. in Economics, Associate Professor, Department of Economics and Management in the Agro-Industrial Complex, Orel State Agrarian University named after N.V. Parahin; 69 General Rodina Str., Oryol, 302019, Russia; e-mail: satynter@mail.ru.

Nikolay Aleksandrovich YAKOVLEV: Ph.D. in Agriculture, Associate Professor, Department of Economics and Management in the Agro-Industrial Complex, Orel State Agrarian University named after N.V. Parahin; 69 General Rodina Str., Oryol, 302019, Russia.

Galina Petrovna ZVEREVA: Ph.D. in Economics, Associate Professor, Department of Economics and Management in the Agro-Industrial Complex, Orel State Agrarian University named after N.V. Parahin; 69 General Rodina Str., Oryol, 302019, Russia.

Anastasiya Sergeevna VOLCHENKOVA: Ph.D. in Economics, Associate Professor, Department of Economics and Management in the Agro-Industrial Complex, Orel State Agrarian University named after N.V. Parahin; 69 General Rodina Str., Oryol, 302019, Russia.