Scientific Papers Series Management, Economic Engineering in Agriculture and Rural Development
Vol. 22, Issue 1, 2022
PRINT ISSN 2284-7995, E-ISSN 2285-3952

STATISTICAL ASSESSMENT OF THE AVERAGE YIELDS OF CUCUMBERS AND GHERKINS IN SOME BALKAN COUNTRIES

Delyana DIMOVA

Agricultural University - Plovdiv, 12 Mendeleev Blvd, Plovdiv 4000, Bulgaria; E-mail: delyanadimova@abv.bg

Corresponding author: delyanadimova@abv.bg

Abstract

Data concerning harvested area and production of cucumbers and gherkins have been extracted from the website of FAOSTAT. They have been presented and saved in a separate Excel file. The calculated average yields of the indicated crops in five Balkan countries (Greece, Romania, North Macedonia, Albania and Bulgaria) have been studied in the period 1999-2020. The mentioned information has been compared, assessed and summarized. Analysis of variance and Tukey's range test have been applied to the considered data about average yields. The results showed that the indicator (harvested area) for these two crops had significantly bigger values for Romania in the interval from 2001 to 2020. At the same time, the values of the other examined indicator (average yields) for Greece were relatively higher during 2014 as well as 2018. A similar situation was observed for Romania in 2004, 2008 and 2018, while for Bulgaria in the two years - 2014 and 2016. This indicator for Albania and North Macedonia increased gradually during the bigger part of the period. The statistical evaluation of the average yields of cucumbers and gherkins for the studied Balkan countries formed three groups with statistically significant differences.

Key words**:** *analysis of variance, assessment, Balkan countries, cucumbers and gherkins*

INTRODUCTION

Due to the fast growth of internet there has been big volume of information is produced and shared by various administrations in nearly every business, industry and other fields [8]. Due to this high explosion it's really a big challenge to store, manage and access knowledge from this [8].

Organisations gather increasingly large and complex data sets each year [7]. After the data is stored in a precise structured format, further analysis and visualization of data is essential to discover the hidden valuable insight from the large dataset (Mahajan & Gokhale, 2019 [13]). The study of Khusainova, R. M., et al., 2016 [9] notes that the application of statistical methods in economics plays an important role, as it deals with the processing and analysis of vast amounts of information about socio-economic phenomena [9].

The information about two vegetable crops has been studied by using some of these statistical methods in the current work. It should be emphasized that growing vegetables in certain regions is particularly important. The study of Schreinemachers, P et al., 2018 [17] notes that now is the time to prioritize investments in vegetables, thus providing increased economic opportunities for smallholder farmers [17]. World vegetables production grew faster between 2000 and 2019 [4]. The share of onions, cucumber and gherkins, and eggplants increased, while that of cabbages almost halved and that of tomatoes remained stable according to FAO 2021 [4]. In 2019, the East European cucumber and gherkin market increased by 9.1% to $4.5B, rising for the third consecutive year after four years of decline (www.globaltrademag.com [6]). In addition, cucumbers and gherkins are a group of cucurbitaceous vegetables that are mainly used as salad vegetables [1]. Specially, data concerning the indicated two crops are the object of consideration in this study.

The aim of this article is to present a statistical assessment of the average yields of cucumbers and gherkins in five Balkan countries (Greece, Romania, Albania, North Macedonia and Bulgaria) during 1999-2020.

MATERIALS AND METHODS

The paper considers two basic characteristics of the listed above crops. In this connection, the indicators - harvested areas and production have been examined for the indicated time interval. The mentioned information has been provided from the website of FAOSTAT [5]. It has been extracted and organized into an Excel file. The average yields of the cucumbers and gherkins for the presented time period in these five Balkan countries have been calculated. Using lists of data and filters in Excel [14], [3] users could display the relevant subsets [10] for the surveyed objects.

The examined information about the indicated crops for Greece, Romania, North Macedonia, Albania and Bulgaria has been compared, assessed and summarized. The statistical method as analysis of variance [16] has been used for studying the average yields of the considered crops. The main task of analysis of variance is to determine the individual or combined influence of one or more indicators (factors) on another indicators and to evaluate these influences [12]. In addition, the Tukey's test [18] is used for the investigated objects in this work.

The percentage change (H_{ij+1}) of the considered indicator (average yields) for each year compared with the preceding one has also been calculated:

$$H_{ij+1} = \frac{h_{i\,j+1}}{h_{ij}} * 100 - 100$$

where: $h_{i\,j+1}$ and $h_{i\,j}$ - the average yields of the studied crops for i^{-th} country during j^{-th} and $j+1^{-st}$ year, $1<=i<=5$, $1<=j<=21$.

The R Commander [15] and MS Excel [11], program products have been used for the statistical data processing [2]. The obtained results have been presented in tabular or graphic form and the relevant conclusions have been drawn.

RESULTS AND DISCUSSIONS

The information concerning the mentioned crops has been studied during 22 years period. Applying certain filters, the following subsets of data are obtained and visualized:

- One or more indicators for the examined time interval in the considered countries;
- Indicators values for the listed objects for one or more years;
-Chosen value for a given indicator (harvested area, production or average yields) in certain country;

As can be seen from Figure 1, the obtained subset about harvested area in the considered Balkan countries for two years time interval is presented. It should be noted that the other examined elements could also be found and displayed for definite interval of years.

	A	B	C	D	E
1	Area	Element	Year	Unit	Value
2	Albania	Area harvested	1999	ha	1300
3	Albania	Area harvested	2000	ha	1100
24	Romania	Area harvested	1999	ha	8336
25	Romania	Area harvested	2000	ha	9031
46	Bulgaria	Area harvested	1999	ha	10351
47	Bulgaria	Area harvested	2000	ha	9495
68	Greece	Area harvested	1999	ha	2819
69	Greece	Area harvested	2000	ha	2819
90	North Macedonia	Area harvested	1999	ha	1183
91	North Macedonia	Area harvested	2000	ha	1197

Fig. 1. Obtained objects for the two years period
Source: Data from FAOSTAT [5].

The analysis of the data on harvested area showed that the values of this surveyed indicator were quite higher for Romania in comparison with those ones for the other four countries in the interval 2001-2020 (Fig. 2). A significant decline of this variable for cucumbers and gherkins in the indicated country was calculated for 2018. In this case, it was more than 45%.

The same process was observed in Bulgaria for the period 1999-2001 and 2003-2005. The decrease of the indicator was about 61% for the first three years and more than 80% during the second three years period.

The situation was quite different for the examined information about harvested area for the listed crops in Albania.

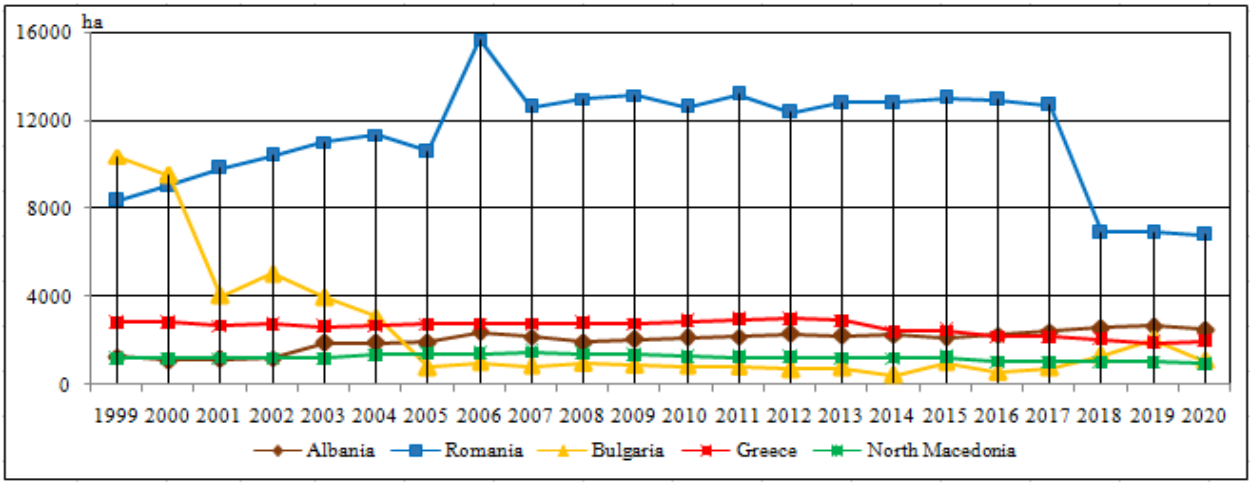

Fig. 2. Graphical analysis of the harvested area for the studied countries during 1999-2020
Source: FAOSTAT [5].

The growth of the indicator was 53.59% during the first seven years (2000-2006) of the interval. The same dependences were established from 2016 to 2019, but in this case the increase was significant smaller about 14.54%.

The reduction of the indicator values for Greece was more than 22.70% in the years between 2015 and 2019. The reverse process was observed in two time intervals. The first of them includes 2003-2007, while the second one contains 2009-2012. The increase of the harvested area in these periods was about 4.5% and 7%, respectively.

The change of the indicator for cucumbers and gherkins in North Macedonia was insignificant during 1999-2002 as well as 2016-2019 (Fig. 2). The period from 2003 to 2007 is characterized with a continuous growth of this studied variable from about 23%. A certain decline was calculated during 2008-2013. It was about 12.86%.

Analyzing the presented results from Figure 3, it can be summarized that the production of cucumbers and gherkins for Romania in 2004, 2011, 2015 and 2017 is quite higher in comparison with this one for the other studied countries. Three periods are formed where the values of this indicator for Romania increased. The first of them includes 2000-2004, while the second and the third one contains the years 2007-2011 and 2013-2015. In this case the growth of the variable was more than 1.9, 1.6 and 1.16 times respectively in the listed subintervals. A quite big decline was observed at the last four years of the considered period. It was about 1.7 times.

The production of cucumbers and gherkins in Albania increased continuously for almost all studied time interval with exception of some individual years – 1999-2000, 2006-2007 as well as 2020. The growth of the variable was more intensive in the years between 2015 and 2019. It was more than 1.5 times.

A similar situation was observed for the production of the examined crops in North Macedonia in the periods 1999-2006, 2008-2012 and 2014-2016. A certain reduction of the indicator was established during the last three years. In addition, the same dependences were observed in 2007 and 2013.

An interesting fact should be noted. Relatively high values of the production of cucumbers and gherkins in Greece were visualized on the diagram from figure 3 for the considered 22 years interval. But a slight decline of this indicator was calculated for five consecutive years from 2010 to 2014. It was about 8%. The lowest value was registered in 2017.

A significant decrease of the production from these crops in Bulgaria was established for the first three years of the surveyed period. It was more than 2.7 times. The same process was observed for the subinterval from 2009 to 2012. It should be noted that the decrease here was smaller (about 2.05 times). The indicator grew slightly in 2002-2004, 2007-2008 as well as 2013-2016.

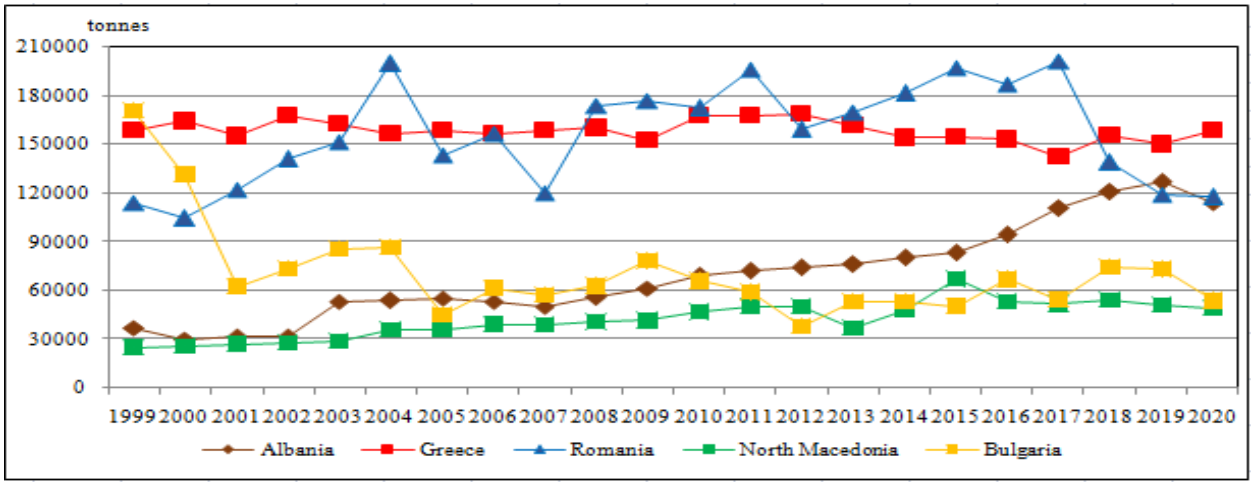

Fig. 3. Visualization of the produced quantities for the studied crops in 22 years interval
Source: FAOSTAT [5].

Another fact is of an interest. The highest value of the examined indicator for Bulgaria was registered for the first year (1999) of the time interval.

The current work evaluates the average yields of cucumbers and gherkins for the listed above five countries in the years between 1999 and 2020. For this purpose, the method of the analysis of variance (Anova) is applied. As can be seen from table 1, the calculated value of the significance level is less than α (α=0.05). Therefore, there are statistically significant differences between the considered data.

Table 1. The obtained results from the performed analysis

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	32,535.61	4	8,133.90	32.11	2E-17	2.458
Within Groups	26,596.96	105	253.30			
Total	59,132.57	109				

Source: Own calculation on the basis of data from FAOSTAT [5].

The results from Tukey's test are displayed on Table 2. They presented the following groups, which were obtained from the performed assessment:

- Bulgaria and Greece are included in one group. The values of the examined indicator (average yields) for these countries are the higher;
- North Macedonia and Albania are presented in the next group. There are no statistically proven differences in the average yields of the listed crops in these two countries;
- The lowest values of average yields of cucumbers and gherkins during the studied time period is observed in Romania. The indicated country forms a separate group.

Table 2. Evaluation of the data concerning the average yields (t/ha)

Considered countries	Evaluation	
Romania	14.142	a
Albania	33.209	b
North Macedonia	35.151	b
Bulgaria	56.380	c
Greece	61.736	c

Values in column followed by the same letter do not differ significantly
Source: Own calculation on the basis of data from FAOSTAT [5].

This work also calculates the percentage change in the average yields of the examined crops for each year compared with the preceding one for the indicated countries. The obtained results from the data processing were presented on Table 3. Relatively higher values of the variable H_{ij+1} (i=1, j=15, j=19) were established for two nonconsecutive years in Greece. Therefore, during 2014 as compared to 2013, the average yields grew by 14.86%. A similar situation was observed for 2018 where the increase was about 15%.

The variable H_{ij+1} (i=2, j=9) was significantly higher for Romania in 2008. Compared to the

year 2007, the average yields of cucumbers and gherkins increased by 40.78%. The same process was observed in 2004 and 2018, where H_{ij+1} (i=2, j=5, j=19) was 28.26% and 26.73%, respectively. A bigger decline was calculated during 2000, 2005-2006 as well as 2019. It has varied in range from approximately 14.41% to 26%.

Table 3. Percentage change (H_{ij+1}) of the average yields for each year of the examined period

Year	Greece	Romania	Albania	North Macedonia	Bulgaria
2000	3.74	-15.79	-4.18	1.77	-16.14
2001	-0.89	7.96	2.03	2.52	13.01
2002	5.62	8.98	-4.17	2.47	-6.29
2003	0.88	1.18	5.20	5.91	46.70
2004	-5.29	28.26	0.38	9.01	30.23
2005	-1.25	-23.35	1.73	-1.40	105.85
2006	-1.81	-26.05	-22.06	8.35	7.71
2007	1.50	-4.67	1.80	-5.54	8.44
2008	-0.06	40.78	27.65	10.15	-3.94
2009	-4.19	0.38	2.99	5.25	37.77
2010	6.33	1.42	6.90	19.94	-13.53
2011	-3.46	9.16	2.46	7.465	-2.740
2012	0.31	-13.56	-1.69	2.13	-27.88
2013	-2.13	2.94	7.13	-24.89	30.54
2014	14.86	6.99	1.76	30.79	79.48
2015	-1.41	6.90	10.03	34.51	-60.37
2016	10.96	-4.87	7.72	-4.86	144.84
2017	-5.55	9.80	8.52	-2.04	-41.07
2018	15.00	26.73	3.56	5.50	-23.30
2019	4.42	-14.42	2.13	-6.25	-35.85
2020	1.58	0.45	-4.66	5.76	42.14

Source: Own calculation on the basis of data from FAOSTAT [5].

The values of the variable H_{ij+1} were positive during almost all surveyed time interval for North Macedonia. They were negative for six nonconsecutive years. The reduction was quite big only for 2013. It was more than 24.89%. In addition, during 2015 as compared to 2014, the average yields of cucumbers and gherkins grew by 34.51%.

A similar situation was observed for the studied values of the variable H_{ij+1} in Albania. The average yields of the indicated crops increased gradually during 2013-2019, 2007-2011, 2003-2005 and 2001. The significant growth was calculated in 2008. Compared to the year 2007, the change of the indicator H_{ij+1} (i=3, j=9) was about 27.65% (Table 3). A big decline of the average yields of cucumbers and gherkins was established in 2006.

The calculations show that the variable H_{ij+1} (i=5, j=6, j=15, j=17) has relatively higher values for the following years in Bulgaria - 2005, 2014 as well as 2016. During 2015 as compared to 2014, the average yields of the mentioned crops decreased by 60.37%. This obtained decline was the biggest during the studied time period.

Summarizing the results from the data processing, it should be noted that an increase in the average yields of these two investigated crops was registered in the following countries - Romania, Greece, North Macedonia and Bulgaria. At the same time a small decline of the indicator was observed only in Albania.

CONCLUSIONS

The studied information related to the indicated crops (cucumbers and gherkins) has been extracted from the website of FAOSTAT. It has been structured and saved in a separate xlsx file. The average yields of the cucumbers and gherkins during the considered time interval in five Balkan countries have been calculated.

The article presents a statistical assessment of the average yields of the indicated crops in these studied countries (Greece, Romania, North Macedonia, Albania and Bulgaria) for the period from 1999 to 2020.

The obtained results from the data processing showed the following:

- The values of the indicator (harvested area) for these two crops were quite higher for Romania in comparison with those ones for the other four countries during 2001-2020. A certain growth of the variable for Albania was calculated in the period 2000-2006 and 2016-2019. The same process was observed in North Macedonia for 2003-2007, as well as in Greece for two subintervals 2003-2007 and 2009-2012. The decrease of the indicator for

Bulgaria was significant for the period 1999-2001 and 2003-2005.

- The values of the second studied indicator (production of the listed crops) for Romania were significantly bigger for the period 2004, 2011, 2015 and 2017, while for Greece they were comparatively high in 1999-2020. One steady increase of the variable was observed for almost all studied time interval in Albania and North Macedonia, but in Bulgaria the reduction of the production of these crops was quite big in the first three years - 1999-2001;
- The average yields of cucumbers and gherkins for Greece were significantly higher in 2014 and 2018. A similar situation was observed for Romania in 2004, 2008 and 2018, while for Bulgaria in two years - 2014 and 2016. This indicator for Albania and North Macedonia increased gradually during the bigger part of the considered time interval;
- The statistical evaluation of the average yields of cucumbers and gherkins for the studied five countries formed three groups with statistically significant differences.

REFERENCES

[1]Agrihortico CPL, 2021,Cucumbers and Gherkins, Agrihortico,https://www.storytel.com/bg/bg/books/cucumbers-and-gherkins-1189399, Accessed on Jan. 14th, 2022.

[2]Dimova, D., 2021, Survey of the Purchasing Power of Households Concerning Certain Milk Products, Scientific Papers Series Management, Economic Engineering in Agriculture and Rural Development, Vol.21(1):233-236.

[3]Dimova, D., 2021, Mathematical Models Describing the Dynamics in Average Prices and Purchased Quantities of Fresh Fruits, Scientific Papers Series Management, Economic Engineering in Agriculture and Rural Development, Vol. 21(2):237-242.

[4]FAO, 2021, World Food and Agriculture – Statistical Yearbook 2021. Rome, https://doi.org/10.4060/cb4477en, https://www.fao.org/3/cb4477en/online/cb4477en.html, Accessed on Jan. 14th, 2022.

[5]FAOSTAT, https://www.fao.org/faostat/en/#data/QCL, Accessed on Dec. 22, 2021.

[6]IndexBox, 2021, Driven by rising demand in Russia and Ukraine, the East-European cucumber and gherkin market to see solid growth, Written by IndexBox, https://www.globaltrademag.com/driven-by-rising-demand-in-russia-and-ukraine-the-east-european-cucumber-and-gherkin-market-to-see-solid-growth/, July 16th, 2020, Accessed on Jan. 14th, 2022.

[7]Kandel, S., Paepcke, A., Hellerstein, J., Heer, J., 2012, Enterprise data analysis and visualization: An interview study, IEEE Transactions on Visualization and Computer Graphics, Vol.18, no. 12: 2917-2926.

[8]Kanimozhi, K. V., Venkatesan, Dr. M., 2015, Unstructured Data Analysis-A Survey, International Journal of Advanced Research in Computer and Communication Engineering, Vol. 4(3):223-225.

[9]Khusainova, R. M., Shilova, Z. V., Curteva, O. V., 2016, Selection of Appropriate Statistical Methods for Research Results Processing, Mathematics Education, 11(1):303-315.

[10]Kuneva, V., Milev, M., Gocheva, M., 2021, Modeling the Transportation Assesment with MS Excel Solver, AIP Conference Proceedings, vol. 2333(1): 150005-1–150005-10.

[11]Levine, D. M., Stephan, D.F., Szabat, K. A., 2016, Statistics for Managers Using Microsoft Excel, 8th Edition, Pearson, USA.

[12] Lidanski, T., 1988, Statistical Methods in Biology and Agriculture, Zemizdat, Sofia, [in Bulgarian].

[13]Mahajan K. N., Gokhale L. A., 2019, Advanced Charting Techniques of Microsoft Excel 2016 Aiming Visualization, International Journal of Computer Sciences and Engineering, India, vol.7(1):198-207.

[14]Mihaylov, D., 2016, Excel 2013, New star, Sofia, [in Bulgarian].

[15]Nguyen-Feng, V., Stellmack, M.A., 2016, A Guide to Data Analysis in R Commander, University of Minnesota, https://manualzz.com/doc/28971240/a-guide-to-data-analysis-in-r-commander--rcmdr-, Accessed on July 16, 2021.

[16] Sawyer, S. F., 2009, Analysis of Variance: The Fundamental Concepts, The Journal of Manual & Manipulative Therapy, volume 17, number 2: E27-E38.

[17]Schreinemachers, P., Simmons, E. B., Wopereis, M. C.S., 2018, Tapping the economic and nutritional power of vegetables, Global Food Security, Vol. 16: 36-45.

[18]Tukey, J. W., 1949, Comparing Individual Means in the Analysis of Variance, International Biometric Society, Biometrics, Vol. 5, No. 2: 99-114.