Czech University of Life Sciences Prague
Faculty of Economics and Management

agris on-line

Papers in Economics and Informatics

http://online.agris.cz

ISSN 1804-1930
XII, 2020, 4

International scientific journal
Prague

Agris on-line Papers in Economics and Informatics

The international reviewed scientific journal issued by the Faculty of Economics and Management of the Czech University of Life Sciences Prague.

The journal publishes original scientific contributions from the area of economics and informatics with focus on agriculture and rural development.

Editorial office

AGRIS on-line Papers in Economics and Informatics
Department of Economics of FEM CZU Prague
Kamýcká 129, 165 00 Praha-Suchdol
Czech Republic
Phone: +420 224 382 056
E-mail: agrisonline(at)pef.czu.cz

Publisher

Faculty of Economics and Management
Czech University of Life Sciences Prague
Kamýcká 129, 165 00 Praha-Suchdol
Czech Republic
Reg. number: 60460709

ISSN 1804-1930

XII, 2020, 4
30th of December 2020
Prague

Agris on-line
Papers in Economics and Informatics

ISSN 1804-1930

XII, 2020, 4

Agris on-line Papers in Economics and Informatics

Volume XII Number 4, 2020

Content:

Volume XII Number 4, 2020

Accounting for TFP Growth in Global Agriculture - a Common-Factor-Approach-Based TFP Estimation

Lajos Baráth, Imre Fertő

Institute of Economics, Centre for Economic and Regional Studies, Budapest,Hungary

Abstract

There is no consensus about trends in agricultural productivity among agricultural economists. The aim of this paper is to contribute to the investigation of this issue by estimating a Total Factor Productivity (TFP) index for global agriculture and global agricultural regions. One of the biggest challenges with analysing global productivity trends is the lack of price data or cost shares, especially in developing countries. We apply recently introduced econometric models that permit accounting for technology heterogeneity and the time-series properties of data to estimate cost shares. Aggregate sectoral data from the USDA ERS database are investigated for the period 1990 to 2013. Although we used a different method, our results are in line with earlier findings that used USDA or FAO database. TFP growth has accelerated in world agriculture, largely due to better performance in transition countries. Although TFP growth has accelerated in world agriculture, it has slowed down in industrialized countries. TFP growth in the EU has increased, but at slower rate in recent years. In the Old Member States the growth rate has decreased, whereas in the New Member States it has increased. The results highlight that insufficient spending on productivity-enhancing agricultural R&D in industrialized countries may put future agricultural productivity growth at risk.

Keywords

Total Factor Productivity (TFP), agricultural productivity, heterogeneous technology, time series properties, cross sectional dependence.

Baráth, L. and Fertö, I. (2020) "Accounting for TFP Growth in Global Agriculture - a Common-Factor-Approach-Based TFP Estimation", *AGRIS on-line Papers in Economics and Informatics*, Vol. 12, No. 4, pp. 3-13. ISSN 1804-1930. DOI 10.7160/aol.2020.120401.

Introduction

Although much work has been done on analysing trends in agricultural productivity, understanding of this issue remains far from complete. Even among agricultural economists that study productivity, there is no consensus about whether the rate of growth in agricultural productivity is slowing (Fuglie and Wang, 2012). Recently, Alston, Babcock, and Pardey (2010) examined a number of studies about trends in agricultural productivity in various regions of the world. Their conclusion was that "agricultural productivity has slowed, especially in the world's richest countries." But they also recognized that the evidence was mixed, and, given the importance of the issue, that it needed further investigation (Fuglie and Wang, 2012).

The goal of this paper is to contribute to the investigation of this issue. We use recently introduced methodological developments to provide insight into the questions: (i) whether global agricultural TFP growth has slowed down in recent decades, and (ii) whether the slowdown in productivity is more significant in industrialized countries. Additionally, we investigate the evolution of TFP growth in the European Union and examine the differences in TFP growth between Old and New Member States.

TFP is usually defined as the ratio of aggregate output to aggregate input. It is therefore necessary to account for the sum of changes of outputs and inputs used in production. We apply the 'growth accounting' method.

The growth accounting method measures aggregate input growth as the weighted sum of the growth rates of the quantities of the individual factors of production, wherein the weights are the cost shares. In the case of outputs, revenue shares are used. However, for most countries in the world there is a lack of representative data about input prices and therefore cost shares. This is especially true for developing countries, where the most important inputs are farm-supplied, like land and labour, but where wage labour and land rental markets are thin, making it difficult to assess the share of these inputs in total cost (Alston, Babcock and Pardey, 2010; Fuglie, 2012).

To deal with this challenge, most examinations of global agricultural TFP have relied on distance function measures (like the Malmquist index) to compare productivity among groups of countries. Distance functions are derived from input-output relationships based on quantity data only: for example, Ludena (2010) and Coelli and Rao (2005) applied this method. However, this methodology is sensitive to the set of countries that is included for comparison and the number of variables in the model, and the dimensionality issue (Alston, Babcock and Pardey 2010; Fuglie, 2012; Lusigi and Thirtle, 1997).

Another way of dealing with the lack of input prices was proposed by Avila and Evenson who used input cost shares estimated from agricultural censuses in Brazil and India to impute cost shares for other developing countries. In contrast to many DEA models which found that agricultural TFP growth was negative, the former authors reported positive and accelerating TFP growth for developing countries (Fuglie, 2015; Dias Avila and Evenson, 2010).

Alternatively, econometric estimates of a production function can be used instead of price or cost data. One of the disadvantages of this approach is that it involves strong technical and economic assumptions, like profit maximization and the imposition of a functional form. However, Fuglie argues that imposing more structure could be an advantage when dealing with data with a high degree of measurement error, as it can help to produce more plausible results (Fuglie, 2012; Nin-Pratt et al., 2015). One of the central focal points of studies that used econometric estimation of empirical cross-country production function over the past two decades has been the endogeneity of inputs and, closely related, potential reverse causality in the estimation equation. In the literature, identification in the face of these difficulties is typically achieved through instrumentation (Eberhardt and Teal, 2017). However, if technology is heterogeneous across countries then none of the standard instrumentation strategies applied in the cross-country empirical literature (such as instrumentation using z variables or lags) are valid, since the empirical specifications in these cases assume technology homogeneity (Eberhardt, 2009). In addition, standard instrumentation strategies also assume stationary variable series, as well as cross-sectional independence, and this identification strategy is invalid if any of these assumptions are violated (Eberhardt and Teal, 2013a).

We use advances from non-stationary econometrics and apply a common factor framework to model production. This approach is able to account for heterogeneity in technology, the non-stationary, cross-section dependence of data, and endogeneity.

We follow Eberhardt and Teal (2013b) and compare different heterogeneous models with different assumptions concerning technology heterogeneity and the effect of common factors. We base our decision concerning the preferred model on residual diagnostic tests, and we check for non-stationarity and cross-sectional independence of the residuals.

Although the models thus applied can account for endogeneity, we cannot rule out reverse causality. In order to address this issue, we follow Eberhardt-Teal and simply also estimate the Fully Modified Ordinary Least Square (FMOLS) version of the preferred model and compare estimates between the OLS and FMOLS versions of the models (Eberhardt and Teal 2017). Since the FMOLS methodology is robust to reverse causality, this supplies the assurance that if the coefficients are similar in the two versions, then our estimates represent production function coefficients and our model is not misspecified (e.g., not investment or labour demand equations).

In the second step of our analysis we use the parameter estimates of the preferred model to construct the TFP index and answer our empirical questions. We use the USDA-ERS agricultural database, which has a sufficient number of cross-sectional and time-series observations to model production through applying a common factor framework. Moreover, to the best of our knowledge this database has not yet been examined using these types of models.

Materials and methods

We adopt common factor representation for a production function which allows for heterogeneity in technology, as well as for common shocks to production and/or technology spill overs between countries (‘cross-sectional dependence’) The common factor model framework is arguably ideally suited to the analysis of cross-country productivity (Bai, 2009; Chudik, Pesaran and Tosetti, 2011) but has thus far not been applied very widely

(Cavalcanti, Mohaddes and Raissi, 2011; Eberhardt and Teal, 2013a; Eberhardt and Teal, 2013b).

We follow Eberhardt and Teal (2013a) and we modell production in country i at time t, for $i = 1, \dots, N$, $t = 1, \dots, T$ and $m = 1, \dots, k$ as follows:

$$y_{it} = \beta_i' \boldsymbol{x}_{it} + u_{it}\,,\;\; u_{it} = \alpha_i + \boldsymbol{\lambda}_i' \boldsymbol{f}_t + \epsilon_{it} \qquad (1)$$

$$x_{mit} = \pi_{mi} + \boldsymbol{\delta}'_{mi} \boldsymbol{g}_{mt} + \rho_{1mi} f_{1mt} + \dots + \rho_{nmi} f_{nmt} + v_{mit} \qquad (2)$$

$$\boldsymbol{f}_t = \boldsymbol{\varrho}' \boldsymbol{f}_{t-1} + \epsilon_t \text{ and } \boldsymbol{g}_t = \boldsymbol{k}' \boldsymbol{g}_{t-1} + \epsilon_t \qquad (3)$$

This technique has been shown to be extremely powerful and can provide consistent estimates of β_i' or its cross-country average, even if factors are non-stationary, if there are structural breaks in the factors, or whether there is cointegration or non-cointegration between the model variables (Eberhardt and Teal, 2017).

We follow the existing literature and include proxies for labour, agricultural capital, livestock, fertilizer, and land under cultivation as the m observed inputs x_{it} in the model for observed output y_{it} (all variables in logarithms). As Equation 1 shows, u_{it} is represented by a combination of country-specific fixed effects α_i and a set of common factors f_t with factor loadings that can differ across countries (λ_i). Equation (3) specifies the evolution of the common factors and includes the potential for non-stationary factors ($\varrho = 1$, $k = 1$) and thus non-stationary inputs and outputs (Eberhardt and Teal, 2013a).

Some of the unobserved common factors driving the variation in y_{it} in Equation (1) also drive the regressors in (2). This setup induces endogeneity in that the regressors are correlated with the unobservables in the production function equation (u_{it}), making it difficult to identify β_i separately from λ_i and ρ_i (Kapetanios, Pesaran and Yamagata, 2011; Eberhardt and Teal, 2017). In the literature, identification in the face of these difficulties is typically argued to be achieved through instrumentation. However, if any of the assumptions of homogeneous technology, stationary variable series, or cross-sectional independence are violated, the identification strategy through instrumentation may be deemed invalid (Eberhardt and Teal, 2017). In the common factor framework, the resulting endogeneity problem can be tackled by accounting for the presence of the unobservables in the empirical specification (Eberhardt and Teal, 2017; Pesaran and Smith, 1995). In addition, by using diagnostic tests it is possible to check whether the endogeneity concern has been addressed: "By investigating whether residual series are cross-sectionally correlated we can highlight to what extent we have been able to deal with the dependence caused by the unobservable factors and thus indirectly whether we have addressed the endogeneity concern: if residuals are white noise we know that empirical results do not suffer from endogeneity bias specification" (Eberhardt and Vollrath, 2018).

In the empirical section of this paper we employ and compare different heterogeneous models:

(1) Pesaran and Smith (1995) mean group (MG), (2) the heterogeneous version of the CCE estimators (CCEMG), and (3) the Augmented Mean Group Estimator (AMG) (Eberhardt-Teal, 2013a).

All of the employed models make different assumptions regarding β_i, λ_i, α_i, as well as regarding the persistence of the underlying common factors in equations (1)-(3). A detailed description of these models can be found in many papers, thus for reasons of brevity we direct interested readers to Eberhardt (2009); Eberhardt and Teal (2013); Eberhadt-Teal (2017); Pesaran (2006), and Pesaran and Smith (1995).

If the aim were only to estimate some form of average agricultural technology, the mean of the estimated β_i values across all countries could be used. Alternatively, we could look at the average value of β_i-s across sub-groups of countries. Averaging across alternative groups enables us to identify the central tendencies in technology parameters. However, it is important to note that country-specific parameter coefficients should not be viewed in isolation (Pedroni, 2007) because they frequently yield economically implausible magnitudes (Boyd and Smith, 2002; Eberhardt and Teal, 2013a). In other words, each estimate is a noisy signal of the true parameter value, and averaging across groups of countries boosts this signal and reduces noise (Eberhardt and Vollrath 2018). Therefore, we use for further empirical analysis only the averages of the estimated β_i values across all countries and across sub-groups of countries.

In the second step of our analysis we used the averages of the β_i values of the preferred model to calculate the TFP index, similarly to Fuglie (2010; 2015):

$$ln\left(\frac{TFP_t}{TFP_{t-1}}\right) = \sum_i^I R_i \, ln\left(\frac{Y_{i,t}}{Y_{i,t-1}}\right) - \sum_j^J S_j \, ln\left(\frac{X_{j,t}}{X_{j,t-1}}\right) \quad (4)$$

where R_i is the revenue share of the i[th] output and S_j is the cost share of the jth input.

TFP growth is calculated as the difference in aggregate output and the growth in aggregate input. For the empirical examination we used the gross agricultural output from USDA database and aggregated the inputs using the elasticities of the preferred model that was chosen in the first step of our analysis.

One limitation of this method of calculation is that the cost shares are held constant. However, Fuglie (2010) reports with reference to the applied database that there has been movement among the major input categories, but these changes have occurred gradually (over decades). Thus, the likelihood of major biases in productivity measurement over a decade or two is not large. As our aim is to calculate productivity over two decades (1990-2013), the bias in our case is certainly not large.

We employ aggregate sectoral data for agriculture from the USDA ERS database for the period 1990 to 2013[1]. The applied sample represents an unbalanced panel of 173 countries with 24 time-series observations. A detailed description of the variables can be found on the homepage for the USDA ERS database[2].

For output variable (y) we use gross agricultural output measured in international 2005 $. We use five input variables: land, fertilizer, machinery, livestock, and labour.

Land (X_1) represents total agricultural land in hectares of 'rainfed cropland equivalents'[3]. Fertilizer (X_2) represents metric tonnes of N, P_2O_5, and K_2O fertilizer consumption. The livestock variable (X_3) is the total livestock capital on farms in 'cattle equivalents.' Machinery (X_4) is the total stock of farm machinery in '40-CV tractor equivalents'. Labour (X_5) represents the number of economically active adults engaged in agriculture.

[1] Database downloaded in March 2017.

[2] https://www.ers.usda.gov/data-products/international-agricultural-productivity/

[3] This is the sum of rainfed cropland (weight equals 1.00), irrigated cropland (weight varies from 1.00 to 3.00 depending on region) and permanent pasture (weight varies from 0.02 to 0.09 depending on region). https://www.ers.usda.gov/data-products/international-agricultural-productivity/.

Results and discussion

Parameter estimates of the applied models

Table 1 displays the results of estimated models. For all models we report residual diagnostic tests, namely the Pesaran (2007) panel unit root test and the Pesaran (2004) CD test. We use residual diagnostics to choose the preferred empirical models. Further details about the importance of residual diagnostics in empirical modelling can be found in Eberhardt and Teal (2011) and Banerjee and Carrion-i-Silvestre (2015).

All heterogeneous models yield statistically significant technology coefficients, suggesting that the average technology is different among countries.

The estimates of CCEMG and AMG models are similar, whereas the estimated coefficients of the MG model are different. One explanation for this is revealed in simulation studies: for non-stationary and cross-sectionally dependent data, the MG estimates are severely affected by failure to account for cross-sectional dependence (Coakley, Fuertes, and Smith 2006; Eberhardt and Bond, 2009).

All models yield stationary residuals; however, only the AMG model yields both cross-sectionally independent and stationary residuals. This suggests that the AMG model is a better fit for the database.

Although the AMG model is able to account for technology heterogeneity, cross-sectional

	1_MG		2_CCEMG		3_AMG	
	Coef.	P>|t|	Coef.	P>|t|	Coef.	P>|t|
l_land	0.27	0.00	0.19	0.00	0.22	0.00
l_mat	0.02	0.01	0.02	0.01	0.02	0.00
l_cap	0.12	0.04	0.15	0.00	0.10	0.06
l_liv	0.26	0.00	0.18	0.00	0.19	0.00
Stationarity	I(0)		I(0)		I(0)	
CD	3.4 (0.001)		2.52(0.012)		1.32(0.187)	

Note: I(1) stands for stationary residual; I(0) represents non-stationary residual; CD shows the Pesaran (2004) CD statistic, and in brackets the p-value; H0: cross-sectionally independent residual
Source: Authors' calculations

Table 1: Parameter estimates and residual diagnostics of heterogeneous models.

dependence non-stationarity and endogeneity, we cannot rule out reverse causality. To address this issue, we estimated the (FMOLS) version of this model and compared the resulting estimates with the OLS-based version (Table 2). FMOLS methodology is robust to reverse causality: if the coefficients are similar in the two versions then we can rule out the issue of reverse causality and can be sure that our estimates represent production function coefficients. Results of the comparison of OLS- and FMOLS-based estimates are very similar, thus we used the estimates of the AMG model for further empirical analysis.

Variable	Coefficient	t-Statistic
l_land	0.20	30.96
l_mat	0.02	8.58
l_cap	0.12	13.86
l_liv	0.19	36.20

Note: Model was estimated in RATS
Source: Authors' calculations

Table 2: AMG Model using FMOLS.

TFP growth in industrialized, developing, and transition countries

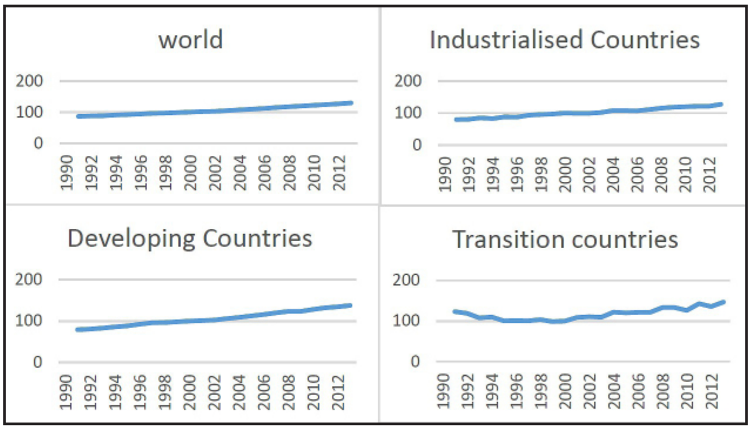
As Fuglie (2010) reported, recent assessments of the global agricultural economy have expressed concern about a significant slow-down in productivity growth. Yet, evidence from major developing countries suggests that productivity growth has accelerated in these regions. This contrasts with the findings of earlier studies of global productivity growth, which found agricultural land and labour productivity rising faster in developed than in developing countries (Hayami and Ruttan, 1985; Craig, Pardey and Roseboom, 1997). Another confounding factor is the uneven performance of agriculture in transition countries. Thus, national and regional evidence is mixed concerning recent trends in agricultural productivity.

The results of more recent papers are also contradictory. Alston and Pardey (2014) find that the global rate of agricultural productivity growth is declining, whereas Fuglie (2015) reports that there has been significant acceleration in global agricultural productivity growth since the 1990s.

Our results are shown in Figure 1. and Table 3. Figure 1 shows the evolution of TFP growth (2000=100%) over the period under analysis, and Table 4 shows the difference in the average annual growth rates for two periods: 1991-2000, and 2011-2012.

Figure 1 shows that TFP growth has increased in global agriculture (world), industrialized (IND), developing (DEV) and transition countries (TRA) compared to 2000 (Figure 1). However, there are remarkable differences in the average annual growth rates in the analysed periods (Table 3).

Examination of average annual TFP growth rates prior and post-2000 shows that TFP growth has



Source: Authors' estimation

Figure 1: Evolution of TFP growth in global agriculture (2000=100%).

Periods	World	IND	TRA	DEV
1991-2000	1.74%	2.58%	-2.21%	2.77%
2001-2013	2.10%	2.07%	2.36%	2.66%

Source: Authors' estimation

Table 3: Average annual growth rate of TFP in global agriculture.

accelerated in world agriculture. It slowed down in industrialized countries, remained nearly at the high-level earlier characteristic of developing countries, and accelerated in transition countries (Table 3).

Although we used different methods to estimate cost shares, our estimates are similar to the USDA estimates (Appendix 1). This suggests that the method we used gives plausible results and is adequate for estimating TFP growth, especially in countries where prices are not available.

These estimates suggest, similarly to earlier findings in the literature, that the acceleration of global TFP growth in recent decades has largely been due to better performance in developing countries and transition economies.

According to Fuglie (2010; 2015), two large developing countries are leading in terms of growth: China, and Brazil, while very recently agricultural TFP growth in India has also accelerated.

In industrialized countries our estimates also confirm earlier findings (Fuglie, 2010; 2015): resources from agriculture are being withdrawn from agriculture at an increasing rate. The average annual growth rate of inputs according to our calculations was -1.2 % from 1991-2000 and was -1.43% from 2001-2013. This calls attention to the same fact that Alston and Pardey (2014) highlight in their paper: insufficient spending on productivity-enhancing agricultural R&D in industrialized countries may put future agricultural productivity growth at risk.

In transition countries at the beginning of 1990s TFP growth slowed down significantly, then in the middle of 1990s started to accelerate. At the beginning of the 2000s the growth rate increased considerably (Figure 1). These results are in line with earlier findings from the literature. Swinnen and Vranken (2010) conducted a detailed examination that involved applying different methods to investigate the changes in productivity in transition countries from 1989-2005. The authors revealed that there have been dramatic changes in productivity over this period in transition countries. In general, one observes a J-shaped (or U-shaped) effect: an initial decline in productivity, and a later recovery. Virtually all countries witnessed an initial decline in productivity, followed by an increase in productivity in the 2000s, and in several transition countries the growth in productivity since 2000 has been quite spectacular (Swinnen and Vranken, 2010). Our graph shows a similar pattern and reveals that in the years following the last year of their examination (2005) productivity also continued to follow this pattern (Figure 1).

TFP growth in the EU

The increase in agricultural productivity has attracted renewed interest in the EU for a number of reasons. First, the European Commission has launched an ambitious program to promote a more resource-efficient Europe by 2020. As a consequence, the agricultural sector is challenged to do more with less. Second, TFP is one of the three impact indicators used in determining the success of the general CAP objective of promoting viable food production. Impact indicators measure the outcome of an intervention beyond its immediate effects. Third, TFP is also used to evaluate the European Innovation Partnership for Agricultural Productivity and Sustainability (EIP-Agri3) (EC, 2016).

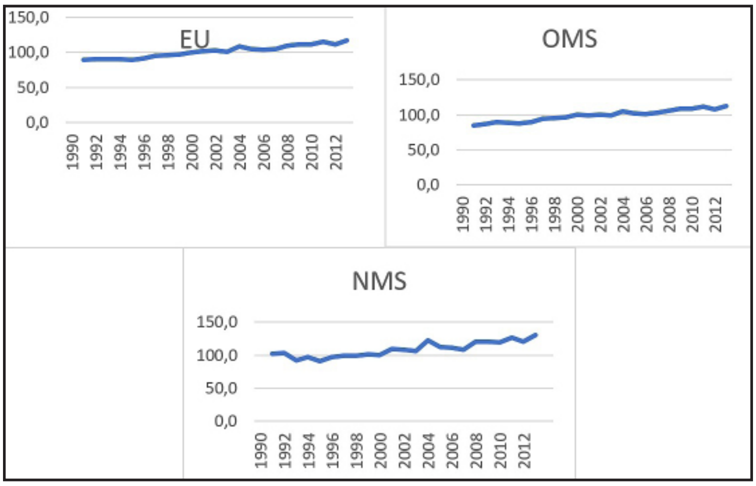
The European Commission in 2016 reported that both in the EU-15 and the EU-N13 TFP growth has increased over the period 1995-2005, and it is remarkable that the high growth rate of the EU-N13 is offset by the lower growth rate in the EU-15; the EU-N13 growth rates are relatively high (over 1.6%/year). In the associated paper the Fisher index was used to estimate TFP, and the Economic Accounts for Agriculture (EAA) database.

In a recent paper, Baráth and Fertő (2017) using the DEA Method constructed a Lowe TFP index based on EAA data and found that TFP slightly decreased in the EU over the period 2004-2013; however, there are significant differences between the OMS and NMS and across Member States.

Our present findings about TFP development in the EU and the OMS and NMS are shown in Figure 2 and Table 4. Figure 2 shows the evolution of TFP growth (2000=100%) while Table 4 shows the average annual growth rates in the 1990s (1991-2000) and 2000s (2000-2013).

Our findings show that TFP growth in the EU has increased over time, although at a slightly slower rate in recent years than in the past. While the growth rate was around 1.2% per year between 1991 and 2000, it had slowed down to around 1.1% between 2001 and 2013.

The differences between the OMS and NMS are remarkable. In the OMS, the growth rate was around 1.7% per year between 1991 and 2000, whereas during this time in the NMS it was only 0.10 %. In the OMS the growth rate slowed down to around 1%; in contrast, in the NMS it was around 1.3 %.



Source: Authors' estimation

Figure 2: Evolution of TFP growth in the EU (2000=100%).

Periods	EU	OMS	NMS
1991-2000	1.22%	1.66%	0.10%
2001-2013	1.11%	1.04%	1.31%

Source: Authors' estimation

Table 4: TFP average annual growth in EU, Oms and NMS.

Direct comparison of these results with those of other studies is difficult because different studies use different groupings and time periods. The EC reports results for the EU-28, EU-15 and EU-N13, whereas the USDA reports results for Europe Northwest, Europe Southern, Europe Transition and Europe Baltic. In the USDA database, 29 European countries can be found, among which 25 EU member states, 14 OMS and 11 NMS. Therefore, we compared our results to the most similar groups (Appendix 2). The comparison shows that in the case of EU estimates our results are similar to those of the USDA, while the EC estimates are much lower. As the EC used the EAAE database, which has more specific variables, the result of our comparison suggests and confirms our earlier finding that FAO database-based analyses likely overestimate productivity growth in the EU (Baráth-Fertő, 2017).

Conclusions

Recent assessments of the global agricultural economy have expressed concern about a significant slowdown in productivity growth (Fuglie, 2010). The first aim of this paper was to examine whether global agricultural productivity has indeed slowed down using recently introduced models which allow us to consider the technological heterogeneity and time-series properties of the data. Our second aim was to examine the differences in TFP growth in global agricultural regions (namely, in industrialized, transition and developing countries), as well as in the EU and its OMS and NMS.

We used diagnostic tests to select the preferred model for further empirical analysis. The results of these showed that the recently introduced AMG model better fits these data.

Our empirical results showed that TFP growth has accelerated in world agriculture over the last two decades. The estimates suggest and confirm earlier findings in the literature that the acceleration of global TFP growth in recent decades has been due to better performance in developing countries and the transition economies (Fuglie, 2010; 2015).

Although we used a different method to the USDA to obtain factor shares, our results are closely in line with their estimates. This suggests that the applied methods give plausible results and can be used to evaluate TFP for global regions and can help to estimate TFP in regions where prices or cost shares are not available.

Our findings show that TFP growth in the EU has increased over time, although at a slightly slower rate in recent years than in the past decade. Differences between the OMS and NMS are remarkable. In the OMS the growth rate has significantly decreased, whereas in the NMS there was a remarkable increase in TFP growth in the last decade. These results are also in line with USDA estimates, but are at odds with EAA-based estimates. This confirms our earlier finding that FAO-data-based analyses likely overestimate productivity growth in the EU (Baráth-Fertő, 2017).

Although TFP growth has accelerated in world agriculture, it has slowed down in industrialized countries. The most important factor determining productivity growth in the long term is innovation, which is driven by research investment. The conceptualization of this process according to Fuglie –Heisey (2007) is as follows: expenditures on agricultural research generate new knowledge that eventually leads to improved technology that is adopted by farmers and technology adoption increases average productivity.

Most studies find a significant positive effect for the productivity of investment in innovative technologies (EC, 2016). Therefore, for stopping or reversing the slowdown in TFP growth in industrialized countries, sufficient spending on agricultural R&D is essential. Additionally, political instruments can increase or decrease TFP growth. However, the link between single political instruments and productivity is not clear; the results of related studies are mixed, especially in the case of agricultural subsidies. Further research which increases understanding of the channels through which agricultural policy instruments affect productivity is also important for improving productivity growth.

Acknowledgments

Lajos Baráth gratefully acknowledges support from National Research Development and Innovation Office (Hungary); project no. K120326.

Corresponding authors
Lajos Baráth
Institute of Economics, Centre for Economic and Regional Studies,
Tóth Kálmán u. 4, 1097 Budapest, Hungary
Phone: +36 30 473 03 12, E-mail: barath.lajos@krtk.mta.hu
ORCID: https://orcid.org/0000-0002-7137-2376

References

[1] Alston, J. M., Babcock, B. A. and Pardey, P. G. (2010) "The Shifting Patterns of Agricultural Production and Productivity Worldwide", The Midwest Agribusiness Trade Research and Information Center, Iowa State UNiversity, Ames, Iowa. ISBN 978-0-9624121-8-9.

[2] Alston, J. M., and Pardey, P. G. (2014) “Agriculture in the Global Economy”, *Journal of Economic Perspectives*, Vol. 28, Bo. 1, pp. 121-146. ISSN 08953309. DOI 10.1257/jep.28.1.121.

[3] Bai, J. (2009) “Panel Data Models with Interactive Fixed Effects”, *Econometrica*, Vol. 77, No. 4, pp. 1229-1279. E-ISSN 1468-0262. DOI 10.3982/ECTA6135.

[4] Baltagi, B. H., Bresson, G. and Pirotte, A. (2007) “Panel Unit Root Tests and Spatial Dependence”, *Journal of Applied Econometrics*, Vol. 22, No. 2, pp. 339-360. E-ISSN 1099-1255. DOI 10.1002/jae.950.

[5] Banerjee, A. and Carrion-i-Silvestre, J. L. (2015) “Cointegration in Panel Data with Structural Breaks and Cross-Section Dependence”, *Journal of Applied Econometrics*, Vol. 30, No. 1, pp. 1-23. E-ISSN 1099-1255. DOI 10.1002/jae.2348.

[6] Baráth, L. and Fertő, I. (2017) “Productivity and Convergence in European Agriculture”, *Journal of Agricultural Economics*, Vol. 69, No. 1, pp. 228-248. E-ISSN 1477-9552. DOI 10.1111/1477-9552.12157.

[7] Cavalcanti, V. de V. , T., Mohaddes, K. and Raissi, M. (2011) “Growth, Development and Natural Resources: New Evidence Using a Heterogeneous Panel Analysis”, *Quarterly Review of Economics and Finance,* Vol. 51, No. 4, pp. 305-318. ISSN 1062-9769. DOI 10.1016/j.qref.2011.07.007.

[8] Chudik, A., Pesaran, M. H. and Tosetti, E. (2011) “Weak and Strong Cross-Section Dependence and Estimation of Large Panels”, *Econometrics Journal*, Vol. 14, No. 1. E-ISSN 1368-423X, ISSN 1368-4221. DOI 10.1111/j.1368-423X.2010.00330.x.

[9] Coakley, J., Fuertes, A.-M. and Smith, R. (2006) “Unobserved Heterogeneity in Panel Time Series Models”, *Computational Statistics & Data Analysis*, Vol. 0, No. 9, pp. 2361-2380. ISSN 0167-9473. DOI 10.1016/j.csda.2004.12.015.

[10] Coelli, T. J. and Rao, D. S. P. (2005) “Total Factor Productivity Growth in Agriculture: A Malmquist Index Analysis of 93 Countries, 1980-2000”, *Agricultural Economics*, Vol. 32, No. 1, pp. 115-134. E-ISSN. DOI 10.1111/j.0169-5150.2004.00018.x.

[11] Craig, B. J., Pardey, P. G. and Roseboom, J. (1997) “International Productivity Patterns: Accounting for Input Quality, Infrastructure, and Research”, *American Journal of Agricultural Economics*, Vol. 79, No. 4. pp. 1064-1076. E-ISSN 1467-8276. DOI 10.2307/1244264.

[12] Dias, A., Flavio, A. and Evenson, R. E. (2010) “Chapter 72 Total Factor Productivity Growth in Agriculture. The Role of Technological Capital”, *Handbook of Agricultural Economics,* Vol. 4, pp. 3769-3822. ISSN 1574-0072. DOI 10.1016/S1574-0072(09)04072-9.

[13] Eberhardt, M. and Bond, S. (2009) “Cross-Section Dependence in non-stationary Panel Models: A Novel Estimator”, Social Research, MPRA Paper,University Library of Munich, Germany.

[14] Eberhardt, M. and Teal, F. (2011) “Econometrics for Grumblers: A New Look at the Literature on Cross-Country Growth Empirics”, *Journal of Economic Surveys*, Vol. 25, No. 1, pp. 109-155. E-ISSN 1467-6419. DOI 10.1111/j.1467-6419.2010.00624.x.

[15] Eberhardt, M. and Teal, F. (2013a) “No Mangoes in the Tundra: Spatial Heterogeneity in Agricultural Productivity Analysis”, *Oxford Bulletin of Economics and Statistics*, Vol. 75, No. 6, pp. 914-939. E-ISSN 1468-0084. DOI 10.1111/j.1468-0084.2012.00720.x.

[16] Eberhardt, M. and Teal, F (2013b) “Structural Change and Cross-Country Growth Empirics”, *World Bank Economic Review*, Vol. 27, No. 2, pp. 229–271. E-ISSN 1564-698X, ISSN 0258-6770. DOI 10.1093/wber/lhs020.

[17] Eberhardt, M. and Vollrath, D. (2018) “The Effect of Agricultural Technology on the Speed of Development”, *World Development*, Vol. 109, pp. 483-496. ISSN 0305-750X. DOI 10.1016/j.worlddev.2016.03.017.

[18] Fuglie, K. and Wang, S. L. (2012) “Productivity Growth in Global Agriculture”, *Population and Development Review*, Vol. 39, No. 2, pp. 361-365. ISSN 00987921.

[19] Fuglie, K. (2015) “Accounting for Growth in Global Agriculture”, *Bio-Based and Applied Economics*., Vol. 4, No. 3, pp. 201-234. E-ISSN 2280-6172, ISSN 2280-6180. DOI 10.13128/BAE-17151.

[20] Fuglie, K. (2012) "Productivity Growth and Technology Capital in the Global Agricultural Economy", In "*Productivity Growth in Agriculture: An International Perspective*", CABI e-Book, Chater No. 16, 333 p. ISBN 9781845939212. DOI 10.1079/9781845939212.0335.

[21] Fuglie, K. and Heisey, P. W . (2007) "Economic Returns to Public Agricultural Research", Economic Brief, Nr. 10, USDA Economic Research Service.

[22] Jerzmanowski, M. (2007) “Total Factor Productivity Differences: Appropriate Technology vs. Efficiency”, *European Economic Review,* Vol 51, No. 8, pp. 2080-2110. ISSN 0014-2921. DOI 10.1016/j.euroecorev.2006.12.005.

[23] Kapetanios, G., Pesaran, M. H. and Yamagata, T. (2011) “Panels with Non-Stationary Multifactor Error Structures”, *Journal of Econometrics*, Vol. 160, No. 2, pp. 326-348. ISSN 0304-4076. DOI 10.1016/j.jeconom.2010.10.001.

[24] Ludena, C. E. (2010) “Agricultural Productivity Growth, Efficiency Change and Technical Progress in Latin America and the Caribbean”, Inter-American Development Bank. DOI 10.2139/ssrn.1817296.

[25] Lusigi, A., and Thirtle, C. (1997) “Total Factor Productivity and the Effects of R&d in African Agriculture”, *Journal of International Development*, Vol. 9, No. 4, pp. 529-538. E-ISSN 1099-1328. DOI 10.1002/(SICI)1099-1328(199706)9:4<529::AID-JID462>3.0.CO;2-U.

[26] Maddala, G. S., and Wu, S. (1999) "A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test", O*xford Bulletin of Economics and Statistics*, Vol. 61, No. S1, pp. 631-652. E-ISSN 1468-0084. DOI 10.1111/1468-0084.0610s1631.

[27] Mundlak, Y., Butzer, R. and Larson, D. F. (2012) "Heterogeneous Technology and Panel Data: The Case of the Agricultural Production Function", *Journal of Development Economics*, Vol. 99, No. 1, pp. 139-149. ISSN 0304-3878. DOI 10.1016/j.jdeveco.2011.11.003.

[28] Nin-Pratt, A., Falconi, C. Ludena, C. L. and Martel, P. (2015) "Productivity and the Performance of Agriculture in Latin America and the Caribbean", Inter-American Development Bank.

[29] Pedroni, P. (2007) "Social Capital, Barriers to Production and Capital Shares: Implications for the Importance of Parameter Heterogeneity from a non-stationary Panel Approach", *Journal of Applied Econometrics*, Vol. 22, No. 2, pp. 429-451. E-ISSN 1099-1255. DOI 10.1002/jae.948.

[30] Pesaran, M. H. (20077) "A Simple Panel Unit Root Test in the Presence of Cross-Section Dependence", *Journal of Applied Econometrics*, Vol. 22, No. 2, pp. 265-312. E-ISSN 1099-1255. DOI 10.1002/jae.951.

[31] Pesaran, M. H. and Smith, R. (1995) "Estimating Long-Run Relationships from Dynamic Heterogeneous Panels", *Journal of Econometrics*, Vol. 68, No. 1, pp. 79-113. ISSN 0304-4076. DOI 10.1016/0304-4076(94)01644-F.

[32] Pesaran, M. H., Smith, L. V. and Yamagata, T. (2013) "Panel Unit Root Tests in the Presence of a Multifactor Error Structure", *Journal of Econometrics*, Vol. 175, No. 2, pp. 94-115. ISSN 0304-4076. DOI 10.1016/j.jeconom.2013.02.001.

[33] Pesaran, M. H. (2004) "General Diagnostic Tests for Cross Section Dependence in Panels General Diagnostic Tests for Cross Section Dependence in Panels", Cambridge Working Papers in Economics 0435, Faculty of Economics, University of Cambridge.

[34] Swinnen, J. and Vranken, L. (2010) "Reforms and Agricultural Productivity in Central and Eastern Europe and the Former Soviet Republics: 1989-2005", *Journal of Productivity Analysis*, Vol. 33, pp. 241–258. E-ISSN 1573-0441, ISSN 0895-562X. DOI 10.1007/s11123-009-0162-6.

Appendix

	World	IND	TRA	DEV
Own estimates				
1991-2000	0.017372	0.025805	-0.00993	0.027712
2001-2013	0.020957	0.02073	0.018062	0.026607
USDA estimates				
1991-2000	0.015999	0.020156	-0.00182	0.022024
2001-2013	0.017261	0.020249	0.014529	0.019734

Source: own processing

Appendix 1: Comparison of own estimates with USDA estimates.

<table>
<tr><th colspan="3">1995-2005</th><th colspan="3">2005-2015</th></tr>
<tr><td rowspan="3">EU-15</td><td>EC</td><td>1.3</td><td rowspan="3">EU-15</td><td>EC</td><td>0.60%</td></tr>
<tr><td>own</td><td>1.5</td><td>own[1]</td><td>1.31%</td></tr>
<tr><td>USDA</td><td>1.48</td><td>USDA[2]</td><td>2%</td></tr>
<tr><td rowspan="6" colspan="3"></td><td rowspan="3">EU-N13</td><td>EC</td><td>1.60%</td></tr>
<tr><td>own[3]</td><td>2.12%</td></tr>
<tr><td>USDA[4]</td><td>2.11%</td></tr>
<tr><td rowspan="3">EU-28</td><td>EC</td><td>0.80%</td></tr>
<tr><td>own[5]</td><td>1.45%</td></tr>
<tr><td>USDA[6]</td><td>2.05</td></tr>
</table>

Notes:
As similar groupings were not available, we compared the estimates of EC, 2016 to the most similar groups as follows:
1: calculated for all Old Member States available in the USDA database
2: calculated as average of Europe Northwest and Europe Southern
3: calculated for all New Member States (countries that joined the EU after May 2004) available in the USDA database
4: calculated as average of Europe Transition and Europe Baltic
5: calculated for all EU Member States available in the USDA database
6: calculated as average of Europe Northwest, Europe Southern, Europe transition and Europe Baltic
Source: own processing

Appendix 1: Comparison of own estimates with USDA estimates.

Improving Agricultural Export Policies in Developing Countries: An Application of Gravity Modelling in the Case of Vietnam's Fishery Export

Manh Hung Do[1], Sang Chul Park[2]

[1] Institute of Development and Agricultural Economics, Leibniz University of Hannover, Germany

[2] College of Political Science and Public Administration, Yeungnam University, Gyeongbuk Province, South Korea

Abstract

This paper investigates the determinants of fishery export from Vietnam using a structural gravity modelling. Taken additional trade-related variables from the World Bank's open data into the estimation of the gravity model, this research will be the first trial to examine the impacts of these variables on export of fishery products and to propose policy implications for stimulating export in Vietnam. The empirical results show that each 1% reduction of export costs might increase approximately 3.7% of the export value of fishery products. This finding is critical because the current administrative system for export of agricultural commodity in Vietnam consists of many stages and includes a long period of animal quarantine inspection, document checking, and customs clearance that might cause additional export costs. Therefore, policies aiming at reducing the costs of border and documentary compliance for export will be significant to stimulate export in developing countries as Vietnam.

Keywords

Fishery export, export policy, free trade agreement, gravity model, random effect estimation, Vietnam.

Do, M. H. and Park, S. C. (2020) "Improving Agricultural Export Policies in Developing Countries: An Application of Gravity Modelling in the Case of Vietnam's Fishery Export", *AGRIS on-line Papers in Economics and Informatics*, Vol. 12, No. 4, pp. 15-27. ISSN 1804-1930. DOI 10.7160/aol.2020.120402.

Introduction

Vietnam's economy has transformed and developed significantly since the political and economic reform in 1986 (Do and Park, 2018) reflected by an annual Gross Domestic Product (GDP) growth rate of about 6.93% between 1987 and 2006 (The Word Bank, 2019). After that period, the country has often been ranked as one of the fastest developing economies at both regional and global levels. In the period of 2007 – 2017, the GDP growth continued to rise significantly at an annual rate of 6.11% that resulted in an increase of the GDP from US$ 77.4 billion in 2007 to US$ 223.8 billion in 2017 (current US$) (The World Bank, 2019).

Apparently, the openness and trade affairs through multilateralism and free trade agreements (FTA) have brought many good opportunities to boost the country's economic development. Vietnam's export value increased from US$ 48.6 billion in 2007 (when it officially became a member of the World Trade Organization) to US$ 213.9 billion in 2017 with a Compound Annual Growth Rate (CAGR) of 16% between 2007 and 2017 (current US$) (ITC, 2019). In other words, trading comprised up to 95% of Vietnam's GDP in 2017. Moreover, this also implies the importance of trading and positive impacts of export stimulation policies on Vietnam's economy.

In 2011, Vietnam targeted itself to become an export-oriented nation with the development of industrial sectors for both agriculture and non-agriculture products by the commencement of the Decision No. 2471/QĐ-TTg dated 28/12/2011 to approve the Strategic policy on export – import between 2011 and 2020, vision to 2030. Although the share of agricultural products in the national export value decreased from 27% to 17% between 2007 and 2017 (MARD, 2019), it is still playing a vital role in providing income sources for nearly 66% of households living in rural regions (Do and Park, 2019; GSO, 2017). Among key sub-sectors in Vietnam's agriculture, fishery export is significantly contributing to national export. The export value of fishery products rose from US$ 3.76 billion in 2007 to about US$ 8.32 billion in 2017 (current US$) which is equal to a CAGR of 8.3% annually (MARD, 2019). Besides, this sub-sector accounted for about 23% of the total export value of agricultural products in 2017 (MARD, 2019). In the Action Plan No. 02/QĐ-BNN-KH dated 02/01/2019 of the Ministry of Agriculture and Rural Development of Vietnam, the fishery sub-sector was aimed to achieve an export value of US$ 10 billion (current US$) in 2020. However, current policies of the fishery sub-sector only focus on production and processing sides with many supporting programs, while the market and export-related aspects have not been paying enough attention in recent years.

In the context of globalization process, Vietnam has achieved some noticeable results. Recently, Vietnam and the European Union (EU) reached a consensus on final texts of the EU - Vietnam Free Trade Agreement (EVFTA) in 2018 and it was successfully signed in Jun 2019. This FTA is critical to Vietnam's economy and its agricultural sector in particular because the EU is one of the key markets of its agricultural products and one of the largest foreign investors in Vietnam (EC, 2019). Particularly, fishery products are among the commodities receiving benefits from the EVFTA for both Vietnam and the EU. With regard to the EU's products, Vietnam will immediately remove all the tariffs at the coming-into-force for fishery products such as salmon, halibut, trout, and rock lobster. The other fishery products from the EU will be liberated after three years. In the case of Vietnam's products, the EU will liberalize its tariffs for non-processed shrimps when the FTA comes into effect and for pangasius and catfish after three years (EU, 2019).

At global levels, fishery sub-sector plays a crucial role in developing countries to provide an important source of livelihood for fish farmers (Allison and Ellis, 2001; Betcherman and Marschke, 2016) and generate income for workers (mostly female labors) in fishery processing industry. Therefore, taking international trade of fishery products to examine its impacts on trade and trade flows is essential to provide evidence for supporting FTAs such as the EVFTA, acquire knowledge of trade determinants of fishery products, and recommend supporting policies on stimulating fishery sub-sector in developing countries.

In the literature of international trade assessment, the gravity model is one of the most popular methods employed to analyze trade flows and impacts of trade since its first introduction by Tinbergen (1962). However, several relevant variables have not been included in trade flow analyses such as the export costs and import costs introduced by the World Bank's open data since 2014. These variables can be scientifically applied as explaining variables of international trade researches. Hence, this paper is aimed at addressing two research questions (i) which key determinants can influence trade between Vietnam and its key importers of fishery products and (ii) what implications can be withdrawn from this research for export stimulation policies in Vietnam.

Literature review

The gravity model has been widely applied for assessments in the field of migration (Backhaus et al., 2015) and, especially, trade flows (Bakucs et al., 2019; Baldwin, 1994; Braha et al., 2017; Cardoso et al., 2017; Hndi et al., 2016; Kepaptsoglou et al., 2010; Maciejewski and Wach, 2019; Wach and Wojciechowski, 2016). The concept of the gravity model was based on the Newton's law of gravity (Shepherd, 2016) indicating that the trade between two economies is affected by their mass and distance. After a thorough review, the authors withdrew two critical notes of gravity modelling application in previous researches with regard to estimation methods and variable selections.

Firstly, there is a diverse application of estimation methods in examining gravity modelling. They have been developed from the Ordinary Least Squares (OLS), to Fixed effect/Random effects, Poisson Pseudo Maximum Likelihood (PPML), and, later, to Poisson Quasi Maximum Likelihood (PQML) in the last two decades. Nevertheless, some scholars mentioned that the OLS used for gravity modelling might contain some methodological and modelling flaws (Kepaptsoglou et al., 2010). Particularly, Anderson and Van Wincoop (2003) and Henderson and Millimet (2008) pointed out that the OLS's implementation assumptions were not consistent with the theoretical models. In addition, Kepaptsoglou et al. (2010) concluded that many empirical researches put more emphasis on the fixed effects approaches because of its appropriateness and the selection of estimation methods would depend on researchers' interests of the analysis, countries' and data's characteristics, and theoretical models.

Secondly, the literature review shows that previous studies only focused on some core explanatory variables such as GDP/GDP per capita/Gross National Income (GNI)/Distance/Population and some additional variables as FTA/common languages/border/exchange rate/tariffs/colonial history in estimating gravity models to explain international flows in the last two decades. However, the variable of distance that is widely used in previous studies is a physical distance that remains unchanged permanently, while trade is dramatically changing in the past few decades because of the globalization process. Therefore, this research proposes an alternative application of new variables to measure distance in gravity modelling that is an average shipping time (e.g. number of days) using sea freight from exporter's international ports to importer's international ports (Table 1). The reason why our research uses this variable is that sea freight shipping is the main measure of transportation for fishery products. Besides, this variable will reflect the practical trading that heavily relies on the development of logistics and shipping services, rather than the static physical distance between the capitals of home and host countries.

Moreover, since 2014, the World Bank's open database has included some additional trade-related variables such as export costs and import costs that could represent practical obstacles to the import and export of a country. Therefore, this paper will be the first trial to apply these variables for the estimation of gravity model to examine determinants of export from Vietnam to its top importers of fishery products (Harmonized System code: 03 and 16) and to propose policy implications for stimulating export.

Materials and methods

Research method

The gravity modelling is a principal measure of scholars who would like to explore and assess the impacts on international trade between countries. In recent years, the gravity models have been widely applied in the field of analyzing trade-related policies with significant improvement of both the uses of variables and estimation methods. In particular, there are many gravity models (such as structural gravity models (Anderson and Van Wincoop, 2003)) applied various fundamental theories in international trade to advance the original mode. These advanced models provide an appropriate platform for conducting simulations of trade impacts and ensure the consistency and unbalanced parameters of estimations (Deardorff, 1998).

No	Dependent variables	Independent variables		Estimation method	Authors
		Core	Additional		
1	Export	GDPs, GDP per capita, population, distance	FTA members, common language, common currency, bilateral exchange rate, political union membership	OLS	Breuss and Egger 1999; Rose, 2000; Feenstra et al., 2001; Sapir, 2001.
2	Export	GDP, GDP per capita, population, distance	Common border, tariffs, common language, country specific factors, remoteness, technological differences	Fixed effect/ random effects	Baltagi et al., 2003; Gopinath and Echeverria, 2004; Cardoso et al., 2017; Maciejewski and Wach, 2019.
3	Imports/ Exports	GDP, distance, population	Land area, common border, island, common language, available trade agreement	Tobit and fixed effects	Soloaga and Winters, 2001.
4	Export/ Import/ Bilateral trade flows	GNI, GDP, distance, population	Common border, tariffs, adjacency, FTA, languages, colony history	PPML and PQML	Siliverstovs and Schumacher, 2008; Lampe, 2008; Braha et al., 2017.

Source: own processing, (Kepaptsoglou et al., 2010, pp. 4-8)

Table 1: Some typical applications of gravity modelling in empirical research.

The intuitive equation of the gravity model from McCallum (1995) is specified as:

$$\log Export_{ij} = c + \beta_1 \log GDP_i + \beta_2 \log GDP_j + \beta_3 \log Distance_{ij} + \varepsilon_{ij} \quad (1)$$

In the Equation (1), $Export_{ij}$ denotes monetary export value from country i to country j, GDPs are each country's (i and j) gross domestic products (economic mass), $Distance_{ij}$ represents an indicator of trade costs which can be the geographical distance between the two countries, and ε_{ij} is the term of random error. Based on the gravity law of Newton, the interpretation of this equation is that larger countries tend to trade more bilaterally and two countries with a larger distance (from each other) will tend to trade less because of higher transportation costs.

However, the intuitive gravity model that was developed in 1960s could not reflect the new advancement of trade literature (Shepherd, 2016). That is why scholars paid more attention to the structural gravity models of Anderson and Van Wincoop (2003) because they include the outward and inward multilateral resistance terms into the trade costs. The structural gravity model (in a short form of aggregate trade) developed by Anderson and Van Wincoop (2003) can be expressed as following:

$$\log Export_{ij} = \log GDP_i + \log GDP_j - \log Y - (1-\sigma)[\log \tau_{ij} - \log \Pi_i - \log P_j] \quad (2a)$$

Whereas,

$$\Pi_i = \sum_{j=1}^{c} \left\{\frac{\tau_{ij}}{P_j}\right\}^{1-\sigma} \frac{Y_j}{Y} \quad (2b)$$

$$P_j = \sum_{i=1}^{c} \left\{\frac{\tau_{ij}}{\Pi_i}\right\}^{1-\sigma} \frac{Y_i}{Y} \quad (2c)$$

$$\log \tau_{ij} = b_1 \log Distance_{ij} + b_2 contig + b_3 comlang\_off + b_4 colony + b_5 comcol \quad (2d)$$

In the equation (2a), the Π_i and P_j are the outward and inward multilateral resistance. The former denotes that exports from country i to country j will rely on the trade costs throughout all export markets. Similarly, the latter indicates that imports of country i from country j will depend on the trade cost from all import markets. This model can reflect an important aspect of international trade that the trade costs of one bilateral flow of export or import might have an impact on all other flows (Shepherd, 2016). Hence, this research paper will employ this structural gravity modelling to assess the trade of fishery products between Vietnam and its key importing partners.

Research data

There are 29 key importers selected to analyze the trade flows of fishery products from Vietnam due to the availability of data (e.g. missing data). These 29 importers account for approximately 80% of total export value from Vietnam between 2014 and 2017 (see Table 5. in the Appendix for the detailed list of selected countries and general information of Vietnam's fishery export). On this list, the Netherlands will represent the EU because this country is one of the largest importers and a major gate of goods from Vietnam to enter other European countries. In other words, the Netherlands is playing an intermediate role in distributing imported products from Vietnam.

Table 2. shows the selected variables and their measurements for the estimation of gravity model. The selection of these variables is mainly relied on previous empirical findings and data's availability. In this regard, the dependent variable is the export value of fishery products from Vietnam and 10 independent variables include Distance from Vietnam to importers; Vietnam's GDP; Importers' GDP; Import costs of importers; Export costs of Vietnam; Ratio of trade to GDP of importers; Ratio of trade to GDP of Vietnam; Foreign direct investment; Tariff levels; and Members of FTA. Among them, the import costs of the importers and export costs of Vietnam are new trial variables in the gravity modelling. These data obtained from the World Bank's open data that are only available since 2014 for the Doing business project.

According to (The World Bank, 2019), the costs of import/export include border compliance and documentary compliance. The former is designed to capture the associated time and cost in order to comply with the country's mandatory regulations for export/import. It also consists of time and cost for operations at ports or borders, customs clearance, and inspection procedures. On the other hand, the latter reflect the time and costs in order to comply with documentary requirements of government agencies in the departure country, destination country, and any transit places. These variables might be practical obstacles to export/import because the high costs of border and documentary compliance might directly affect the export/import process (see Table 3.

Variables	Denotation	Description and Measurement	Data source
Dependent variable			
Fishery Export	EXPORT	Export value of fishery products from Vietnam between 2014 and 2017; US$ Thousand.	(ITC, 2019)
Independent variables			
Distance from Vietnam to importers	DISTANCE	Average time for sea freight shipping from Vietnam's main port to importer's main port; Number of days.	(Linescape, 2019)
Vietnam's GDP	GDP_EXPORTER	Gross Domestic Products of Vietnam between 2014 and 2017; current US$ Billion.	(The World Bank, 2019)
Importers' GDP	GDP_IMPORTER	Gross Domestic Products of Vietnam's fishery importers between 2014 and 2017; current US$ Billion.	(The World Bank, 2019)
Import costs of the importers	COST_IMPORT	Costs associated with the import of the importers including border and documentary compliance; US$ per shipment.	(The World Bank, 2019)
Export costs of Vietnam	COST_EXPORT	Costs associated with the export of Vietnam including border and documentary compliance; US$.	(The World Bank, 2019)
Trade to GDP of importers	OPENNESS_IM	Ratio of Trade to GDP or the economy's openness of the importers; Percentage.	(The World Bank, 2019)
Trade to GDP of Vietnam	OPENNESS_EX	Ratio of Trade to GDP or the economy's openness of Vietnam; Percentage.	(The World Bank, 2019)
Foreign Direct Investment	FDI	Amount of registered FDI capital from the importers in Vietnam; US$ Million.	(MPI, 2019)
Tariff levels	TARIFFS	Average tariff level of fishery products from Vietnam to the importers; Percentage.	(ITC, 2019; (WTO, 2019)
Members of FTA	FTA	Dummy variables; Available bilateral or regional FTA = 1.	(ITC, 2019)

Source: own processing

Table 2: Selected variables and their measurement.

Rank of export value	Country	Export value of fishery products (US$ thousand)	Costs to export (US$)
	World	**169 319 950**	
1	China	22 276 720	569
2	Norway	11 093 251	125
3	Thailand	8 410 162	320
4	**Vietnam**	**8 226 460**	**429**
5	United States of America	7 555 374	235
6	India	7 062 071	474
7	Netherlands	6 119 097	-
8	Canada	5 687 087	323
9	Chile	5 613 779	340
10	Spain	5 378 484	-
11	Germany	5 096 045	390
12	Ecuador	4 471 023	700
13	Sweden	4 245 222	95
14	Indonesia	4 203 170	393
15	Denmark	4 060 645	-
16	Russian Federation	3 650 675	672
17	Poland	3 348 729	-
18	United Kingdom	2 727 897	305
19	Argentina	2 054 631	210
20	Iceland	1 738 431	405

Source: own processing, ITC and the World Bank

Table 3: The world 20 largest exporters of fishery products and their export costs in 2017.

for the summary of the world 20 largest exporters of fishery products and their export costs in 2017). The application of these new variables in the gravity modelling will be the first trial for justifying their significance and impacts on trade. In this research, the authors hypothesize that the variables of import and export costs will have negative impacts on export of Vietnam to the selected countries. In other words, the higher the costs, the lower the export. Hence, a panel data of 29 selected countries importing the vast majority of fishery products from Vietnam between 2014 and 2017 will be used to run the gravity model.

Model specification and estimation method

Based on the theoretical model of Anderson and Van Wincoop (2003), the model specification of this paper is presented as following:

$$\ln EXPORT_{ij} = \alpha_0 + \alpha_1 \ln GDP\_EXPORTER_i + \alpha_2 \ln GDP\_IMPORTER_j + \alpha_3 \ln COST\_EXPORT_i + \alpha_4 \ln COST\_IMPORT_j + \alpha_5 \ln DISTANCE_{ij} + \alpha_6 \ln OPENNESS\_EX_i + \alpha_7 \ln OPENNESS\_IM_j + \alpha_8 \ln FDI_j + \alpha_9 TARIFFS_{ij} + \alpha_{10} FTA_{ij} + \varepsilon_{ij} \quad (3)$$

Whereas,

α_0: is the intercept.

α_1 - α_{10}: are the coefficients of 10 explanatory variables namely *GDP_EXPORTER*, *GDP_IMPORTER*, *COST_EXPORT*, *COST_IMPORT*, *DISTANCE*, *OPENNESS_EX*, *OPENNESS_IM*, *FDI*, *TARIFFS*, and *FTA*.

Among explanatory variables, *GDP_EXPORTER*, *GDP_IMPORTER*, *COST_EXPORT*, *COST_IMPORT*, *DISTANCE*, *OPENNESS_EX*, *OPENNESS_IM*, and *FDI* will be estimated in logarithm.

ε_{ij}: is the random error term.

In the Equation (3), variables such as *GDP_EXPORTER, GDP_IMPORTER, OPENNESS_EX, OPENNESS_IM, FDI*, and *FTA* are expected to have a positive impact on the export value of fishery products, while *COST_EXPORT, COST_IMPORT, DISTANCE*, and *TARIFFS* are supposed to have a negative impact on the export value. It is noted that the panel data that consists of trade data over time could help eliminate biases caused by heterogeneity across observations (Prehn et al., 2016). However, with some time-invariant and relevant variables such as DISTANCE, TARIFFS, and FTA, estimations using fixed effects might cause a perfect collinearity. In this case, random effects would be a more appropriate method for the estimation.

Fundamentally, the fixed and random effects are developed to control over unobserved heterogeneity. However, there are some important differences of the two methodologies. Particularly, the former permit free or structural less variation, whereas the latter require the unobserved heterogeneity to comply with some probability constraints. That is why random effects rely on a strong assumption that unobserved heterogeneity's pattern is randomly distributed to given variance and mean (Gómez-Herrera, 2013). Hence, this research paper will implement the estimation of the gravity model using random effects method. In addition, Breusch and Pagan Lagrangian test for random effects and the Hausman test should be implemented to test whether the random effect model is an appropriate selection.

Results and discussion

Determinants of fishery export from Vietnam

Before conducting the estimation of the gravity model, the authors have implemented the Hausman test and Breusch and Pagan Lagrangian test to justify the appropriateness of the random effect model. The results of the two tests indicate that, under this current specification, the model is firmly fit for the random effect estimation. (see Table 7 and 8 in the Appendix for the detailed results of the Hausman test and Breusch and Pagan Lagrangian test, respectively).

Table 4 shows the empirical results of the gravity model with random effects. Overall, the explanatory variables can explain more than 73% of the response variable by the estimated random effect model and statistically significant at 1%. The results point out that the independent variables are in line with their assumption of expected impacts except for *OPENNESS_EX* and *OPENNESS_IM*. In particular, the variables with positive impacts on the export value consist of *GDP_EXPORTER, GDP_IMPORTER, FDI*, and *FTA*, while the other variables with negative impacts include *COST_EXPORT, COST_IMPORT, DISTANCE*, and *TARIFFS*.

Among the 10 explanatory variables, *GDP_EXPORTER, GDP_IMPORTER, FDI*, especially *COST_EXPORT* and *COST_IMPORT* are statistically significant at or less than 10% level,

Random-effects GLS regression			Number of obs.	=	91
Group variable: pairid			Number of groups	=	25
R-sq:					
within	=	0.3060			
between	=	0.7344			
overall	=	0.7314			
			Wald $chi^2(10)$	=	146.34
corr(u_i, X) = 0 (assumed)			Prob > chi^2	=	0.0000
LNEXPORT	**Coef.**	**Robust Std. Err.**	**P>z**	**[95% Conf.**	**Interval]**
_cons (α_0)	40.6994	20.9698	**0.05**	-0.4006	81.7995
LNGDP_EXPORTER(α_1)	3.3695	2.0537	**0.10**	-0.6556	7.3946
LNGDP_IMPORTER(α_2)	0.6635	0.1335	**0.00**	0.4019	0.9252
LNCOST_EXPORT(α_3)	-3.7203	2.1150	**0.08**	-7.8656	0.4251
LNCOST_IMPORT(α_4)	-0.7206	0.3088	**0.02**	-1.3257	-0.1154
LNDISTANCE(α_5)	-0.2120	0.2199	0.34	-0.6430	0.2190
LNOPENNESS_EX(α_6)	-4.4891	2.9237	0.13	-10.2194	1.2412
LNOPENNESS_IM(α_7)	-0.1502	0.3420	0.66	-0.8206	0.5202
LNFDI(α_8)	0.0370	0.0165	**0.03**	0.0046	0.0694
TARIFFS(α_9)	-0.0028	0.0070	0.69	-0.0164	0.0109
FTA(α_{10})	0.3083	0.3425	0.37	-0.3630	0.9796
sigma_u	0.7437				
sigma_e	0.1988				
rho	0.9333		(fraction of variance due to u_i)		

Source: own processing

Table 4: Empirical results of the gravity model estimation.

while the remaining independent variables are not statistically significant. The authors firmly believe that these insignificant variables are mainly due to a small number of observations. Nevertheless, they still suggest some meaningful interpretations such as *DISTANCE, TARIFFS*, and *FTA*. For example, the coefficient of *DISTANCE* shows that for each percentage of traveling time of sea freight will decrease 0.21% of the export value. The negative impact is also applied for the *TARIFFS* that each 1% increase of the tariff level will relatively reduce 0.0028% value of the export. On the contrary, the coefficient of *FTA* denotes that having a *FTA* between the exporter and importer might help increase 0.31% of the export value.

Regarding the variables with statistical significance, the coefficients of *GDP_EXPORTER* and *GDP_IMPORTER* point out the every 1% increase of these two figures can positively improve the export value of fishery products at 3.37% and 0.67%, respectively. This is understandable with the case of Vietnam since it is an export-oriented country and the fishery sub-sector is contributing significantly to export of agricultural commodities and GDP.

The remaining statistically significant variables including *COST_EXPORT, COST_IMPORT,* and *FDI* imply some remarkable findings. First, the coefficient of *COST_IMPORT* indicates that 1% increase of the import costs from the importers will cause a 0.72% decrease of the export from Vietnam with a statistical significance at 5% level. Nevertheless, this variable is an external variable that is mainly due to conditions and government policies of the importing countries.

Second, the estimation result of *FDI* variable that is statistically significant at 5% level shows that 1% increase of the *FDI* from the importing countries can result in a 0.04% increase of the export value from Vietnam. In other words, the flow of *FDI* capital can positively influence the bilateral trading between the two economies. This is critical because Vietnam is attracting a considerable amount of *FDI* in recent years (MPI, 2019). So, the higher the amount of *FDI*, the higher the trading value between Vietnam and investing countries.

Finally, the coefficient of the *COST_EXPORT* implies that each 1% decrease of the export costs will help to increase approximately 3.7% of the export value of fishery products from Vietnam. This coefficient is statistically significant at 10% level. This finding is critical because it directly relates to Vietnam's administration policies and strategies, as well as the domestic conditions for export such as logistics services and infrastructure. If the country can improve this variable, it might have a significant impact on the export.

Policy implications for Vietnam and developing countries

As mentioned in the research data section, the import and export costs include border compliance and documentary compliance. The former is designed to capture the associated time and cost in order to comply with the country's mandatory regulations for export/import. It also consists of time and cost for operations at ports or borders, customs clearance, and inspection procedures. The latter, on the other hand, reflect the time and cost in order to comply with documentary requirements of government agencies in the departure country, destination country, and any transit places. Since the export costs of Vietnam is remarkably higher than some fishery exporters in the Southeast Asia region such as Thailand and Indonesia (The World Bank, 2019), Vietnam and developing countries should stimulate export through improvements in their administrative system and logistics for export.

Firstly, the administrative system for export of agricultural commodity in Vietnam consists of many stages from animal quarantine (out-port) to customs (in-port) which might include a long period of animal quarantine inspection, document checking, and customs clearance resulting in a higher export cost (Nguyen et al., 2015). In some cases, this might be a cause of corruption since exporters have to bribe officials for a faster inspection or customs clearance. With regard to this aspect, the government should apply more transparent measures such as e-government for its administrative and online declaration in the customs system for export because they might help minimize the paper compliance, reduce conduction time, and increase transparency in export procedures that can save time and money for domestic exporters.

Besides, policies on stimulating export must consider the market sides that can facilitate exports fast, efficiently, and legally. For instance, the government can reduce number of required documents and shorten the implementation period which are currently too costly (in terms of both time and money) in order to comply with documentary requirements of government agencies.

Secondly, the logistic system including infrastructure such as, roads, seaports, and airports in Vietnam and some developing countries is very poor. Exports of agricultural products mainly depend on a few standardized ports in large cities. However, the producing and processing areas are distant from the ports that might take a long period for transportation. This issue is compounded by the poor quality of roads system that might damage quality of products. Therefore, improving logistic system is crucial to reduce the export costs and stimulate export in Vietnam and developing countries.

Conclusion

The results of the gravity model point out some significant findings of the ten explanatory variables. In general, the results show that these variables are having expected impacts (negative/positive) on the dependent variable of fishery export from Vietnam. Among them, five key determinants are statistically significant including GDP of the exporter and importers, FDI capital from the importers into the exporter's territory, and the costs of export and imports.

There is a striking feature that the coefficient of the export cost variable implies that each 1% reduction of the export costs will help increase approximately 3.7% of the export value of fishery products. In the case of Vietnam, it is critical because the current administrative system for export of agricultural commodity in Vietnam consists of many stages and includes a long period of animal quarantine inspection, document checking, and customs clearance that might cause additional export costs. Therefore, this finding is significant because improvements of the government's administrative system can remarkably stimulate exports.

The other five statistically insignificant variables are including openness of the exporter and importers, tariff levels, bilateral or regional free trade agreements, and the distance by sea freight shipping time from the exporter to importers. The reason why these variable are not statistically significant might be due to the data is not large enough and the targeted variables are only available since 2014. Hence, the authors propose a further research applied in a larger number of exporters and importers, as well as a wider range of products in order to justify the impacts of these variables and their statistical significance.

Acknowledgments

The authors would like to thank the Editor and two anonymous reviewers for their valuable comments and suggestions to improve this paper.

Corresponding authors
Sang Chul Park
Department of Public Administration
Yeungnam University
280 Daehak-ro, Gyeongsan, Gyeongbuk 712-749, Republic of Korea
E-mail: scpark@ynu.ac.kr

References

[1] Allison, E. H. and Ellis, F. (2001) "The livelihoods approach and management of small-scale fisheries", *Marine Policy*, Vol. 25, No. 5, pp. 377-388. ISSN 0308-597X. DOI 10.1016/S0308-597X(01)00023-9.

[2] Anderson, J. E. and Van Wincoop, E. (2003) "Gravity with gravitas: a solution to the border puzzle", *American Economic Review*, Vol. 93, No. 1, pp. 170-192. E-ISSN 1944-7981, ISSN 0002-8282. DOI 10.1257/000282803321455214.

[3] Backhaus, A., Martinez-Zarzoso, I., and Muris, C. (2015) "Do climate variations explain bilateral migration? A gravity model analysis", *IZA Journal of Migration*, Vol. 4, No. 1, pp. 1-15. DOI 10.1186/s40176-014-0026-3.

[4] Bakucs, Z., Benedek, Z., and Fertö, I. (2019) "Spatial Price Transmission and Trade in the European Dairy Sector", *AGRIS on-line Papers in Economics and Informatics*, Vol. 11, No. 2, pp. 13-20. ISSN 1804-1930. DOI 10.7160/aol.2019.110202.

[5] Baldwin, R. E. (1994) "*Towards an integrated Europe*", London: Centre for Economic Policy Research, 234 p. ISBN 978-1898128137.

[6] Baltagi, B. H., Egger, P. and Pfaffermayr, M. (2003) "A generalized design for bilateral trade flow models", *Economics Letters*, Vol. 80, No. 3, p. 391-397. ISSN 0165-1765. DOI 10.1016/S0165-1765(03)00115-0.

[7] Betcherman, G. and Marschke, M. (2016) "Coastal livelihoods in transition: How are Vietnamese households responding to changes in the fisheries and in the economy?", *Journal of Rural Studies*, Vol. 45, pp. 24-33. ISSN 0743-0167. DOI 10.1016/j.jrurstud.2016.02.012.

[8] Braha, K., Qineti, A., Cupák, A. and Lazorcáková, E. (2017) "Determinants of Albanian Agricultural Export: The Gravity Model Approach", *AGRIS on-line Papers in Economics and Informatics*, Vol. 9, No. 2, pp. 3-21. ISSN 1804-1930. DOI 10.7160/aol.2017.090201.

[9] Breuss, F. and Egger, P. (1999) "How reliable are estimations of East-West trade potentials based on cross-section gravity analyses?", *Empirica*, Vol. 26, No. 2, pp. 81-94. E-ISSN 1573-6911, ISSN 0340-8744. DOI 10.1023/A:1007011329676.

[10] Cardoso, B. F., Rasetti, M., Giampietri, E., Finco, A. and Shikida, P. F. A. (2017) "Trade Dynamics in the Italian Floriculture Sector within EU Borders: A Gravity Model Analysis", *AGRIS on-line Papers in Economics and Informatics*, Vol. 9, No. 2, pp. 23 - 32. ISSN 1804-1930. DOI 10.7160/aol.2017.090202.

[11] Deardorff, A. (1998) "*Determinants of bilateral trade: does gravity work in a neoclassical world?*", The regionalization of the world economy (pp. 7-32), University of Chicago Press. [Online]. Available: https://www.nber.org/chapters/c7818.pdf [Accessed: 03 Jul. 2019].

[12] Do, M. H. and Park, S. Ch. (2018) "New Rural Development and Hierarchical Governance in Vietnam: Impacts of government support on rural households' income using a Hierarchical Linear Modelling", *AGRIS on-line Papers in Economics and Informatics*, Vol. 10, No. 4, pp. 3-15. ISSN 1804-1930. DOI 10.7160/aol.2018.100401.

[13] Do, M. H. and Park, S. C. (2019) "Impacts of Vietnam's new rural development policy on rural households' income: empirical evidence from the Heckman selection model", *International Review of Public Administration*, pp. 1-17. E-ISSN 2331-7795, ISSN 1229-4659. DOI 10.1080/12294659.2019.1662165.

[14] EC (2019) "*Trade Policy – Vietnam*", The European Commission (EC). [Online]. Available: http://ec.europa.eu/trade/policy/countries-and-regions/countries/vietnam/ [Accessed: 25 Apr. 2019].

[15] EU (2019) "*Guide to the EU – Vietnam Free Trade Agreement*", The European Union (EU). [Online]. Available: http://eeas.europa.eu/archives/delegations/vietnam/documents/eu_vietnam/evfta_guide.pdf [Accessed: 25 Apr. 2019].

[16] Feenstra, R. C., Markusen, J. R. and Rose, A. K. (2001) "Using the gravity equation to differentiate among alternative theories of trade", *Canadian Journal of Economics/Revue canadienne d'économique*, Vol. 34, No. 2, pp. 430-447. E-ISSN 1540-5982. DOI 10.1111/0008-4085.00082.

[17] Gómez-Herrera, E. (2013) "Comparing alternative methods to estimate gravity models of bilateral trade", *Empirical Economics*, Vol. 44, No. 3, pp. 1087-1111. E-ISSN 1435-8921, ISSN 0377-7332. DOI 10.1007/s00181-012-0576-2.

[18] Gopinath, M. and Echeverria, R. (2004) "Does economic development impact the foreign direct investment-trade relationship? A gravity-model approach", *American Journal of Agricultural Economics*, Vol. 86, No. 3, pp. 782-787. E-ISSN 1467-8276. DOI 10.1111/j.0002-9092.2004.00625.x.

[19] GSO (2017) "Main Result from Vietnam Agriculture Census – 2016", General Statistics Office of Vietnam (GSO). [Online]. Available: https://www.gso.gov.vn/defaultaspx?tabid=512&idmid=5&ItemID=18964 [Accessed: 25 Apr. 2019]. (In Vietnamese)

[20] Hndi, B. M., Maitah, M. and Mustofa, J. (2016) "Trade Impacts of Selected Free Trade Agreements on Agriculture: The Case of Selected North African Countries", *AGRIS on-line Papers in Economics and Informatics*, Vol. 8, No. 3, pp. 39 - 50. ISSN 1804-1930. DOI 10.7160/aol.2016.080304.

[21] Henderson, D. J. and Millimet, D. L. (2008) "Is gravity linear?", *Journal of Applied Econometrics*, Vol. 23, No. 2, pp. 137-172. ISSN 1099-1255. DOI 10.1002/jae.974.

[22] ITC (2019) "Trade Map – International Trade Statistics" The International Trade Centre (ITC). [Online]. Available: https://www.trademap.org/tradestat/index.aspx [Accessed: 25 Apr. 2019].

[23] Kepaptsoglou, K., Karlaftis, M. G. and Tsamboulas, D. (2010) "The gravity model specification for modeling international trade flows and free trade agreement effects: a 10-year review of empirical studies", *The Open Economics Journal*, Vol. 3, No. 1. ISSN 1874-9194. DOI 10.2174/1874919401003010001.

[24] Lampe, M. (2008) "*Bilateral trade flows in Europe, 1857–1875: A new dataset*", Research in Economic History, pp. 81-155, Emerald Group Publishing Limited. E-ISBN 978-1-84855-337-8, ISBN 978-1-84855-336-1. DOI 10.1016/S0363-3268(08)26002-7.

[25] Linescape (2019) "*Online platform for finding ocean container shipping routes and schedules*". [Online]. Available: https://www.linescape.com/ [Accessed: 25 Apr. 2019].

[26] Maciejewski, M. and Wach, K. (2019) "What Determines Export Structure in the EU Countries?: the Use of Gravity Model in International Trade Based on the Panel Data for the Years 1995-2015", *Journal of International Studies*, Vol. 12, No. 1, pp. 151-167. E-ISSN 2306-3483, ISSN 2071-8330. DOI 10.14254/2071-8330.2019/12-1/10.

[27] MARD (2019) "*Monthly Statistic Report of Agricultural Sector*", The Ministry of Agriculture and Rural Development of Vietnam (MARD). [Online]. Available: https://www.mard.gov.vn/Pages/bao-cao-thong-ke.aspx# [Accessed: 25 Apr. 2019]. (In Vietnamese).

[28] McCallum, J. (1995) "National borders matter: Canada-US regional trade patterns", *The American Economic Review*, Vol. 85, No. 3, pp. 615-623. E-ISSN 1944-7981, ISSN 00028282.

[29] MPI (2019) "*Statistics of Foreign Direct Investment in Vietnam*", The Ministry of Planning and Investment of Vietnam (MPI). [Online]. Available: http://fia.mpi.gov.vn/ChuyenMuc/172/So-lieu-FDI-hang-thang [Accessed: 25 Apr. 2019]. (In Vietnamese).

[30] Nguyen, T. V., Nguyen, N. C. and Bosch, O. J. (2015) "Applying a systems thinking approach to address the barriers affecting the Vietnamese coffee export industry", *International Journal of Business and Management Review*, Vol. 3, No. 4, pp. 49-65. E-ISSN 2052-6407, ISSN 2052-6393.

[31] Prehn, S., Brümmer, B. and Glauben, T. (2016) "Gravity model estimation: fixed effects vs. random intercept Poisson pseudo-maximum likelihood", *Applied Economics Letters*, Vol. 23, No. 11, pp. 761-764. E-ISSN 1466-4291, ISSN 1350-4851. DOI 10.1080/13504851.2015.1105916.

[32] Rose, A. K. (2000) "Currency unions-one money, one market: the effect of common currencies on trade", *Economic Policy*, Vol. 15, No. 30, pp. 7-45. E-ISSN 1468-0327. ISSN 0266-4658.

[33] Sapir, A. (2001) "Domino effects in Western European regional trade, 1960–1992", *European Journal of Political Economy*, Vol. 17, No. 2, pp. 377-388. ISSN 0176-2680. DOI 10.1016/S0176-2680(01)00032-5.

[34] Shepherd, B. (2016) "*The Gravity Model of International Trade: A User Guide", Economics and Social Commission for Asia and the Pacific (ESCAP)*". [Online]. Available: https://www.unescap.org/resources/gravity-model-international-trade-user-guide-updated-version [Accessed: 25 Jun. 2019].

[35] Siliverstovs, B. and Schumacher, D. (2009) "Estimating gravity equations: to log or not to log?", *Empirical Economics*, Vol. 36, No. 3, pp. 645-669. E-ISSN 1435-8921, ISSN 0377-7332. DOI 10.1007/s00181-008-0217-y.

[36] Soloaga, I. and Wintersb, L. A. (2001) "Regionalism in the nineties: What effect on trade?", *The North American Journal of Economics and Finance,* Vol. 12, No. 1, pp. 1-29. ISSN 1062-9408. DOI 10.1016/S1062-9408(01)00042-0.

[37] Tinbergen, J. J. (1962) "Shaping the world economy; suggestions for an international economic policy". [Online]. Available: hdl.handle.net/1765/16826 [Accessed: 26 Apr. 2019].

[38] The World Bank. (2019) "*World Bank Open Data*" [Online]. Available: https://data.worldbank.org/ [Accessed: 25 Apr. 2019].

[39] Wach, K. and Wojciechowski, L. (2016) "Determinants of inward FDI into Visegrad countries: empirical evidence based on panel data for the years 2000-2012", *Economics and Business Review*, Vol. 2, No. 1, pp. 34-52. E-ISSN 2450-0097, ISSN 2392-1641. DOI 10.18559/ebr.2016.1.3.

[40] WTO (2019) "*Tariff Analysis Online*", The World Trade Organisation (WTO). [Online]. Available: https://www.wto.org/english/tratop_e/tariffs_e/tariff_data_e.htm [Accessed: 25 Apr. 2019].

Appendix

No.	Importer	2014	2015	2016	2017	CAGR of 2014 - 2017 (%)
	World	**7,770,359**	**6,550,718**	**7,039,533**	**8,226,460**	1.9%
	Total of 29 importers	**6,165,205**	**5,219,281**	**5,733,868**	**6,835,164**	3.5%
	Share of 29 importers (%)	**79%**	**80%**	**81%**	**83%**	
1	USA	1,710,337	1,314,867	1,440,634	1,414,157	-6.1%
2	Japan	1,193,543	1,032,723	1,096,409	1,305,326	3.0%
3	China	468,558	447,827	681,888	1,089,861	32.5%
4	Korea, Republic	651,747	572,724	607,906	786,105	6.4%
5	Netherlands	211,049	166,957	203,710	304,726	13.0%
6	Thailand	182,381	216,162	241,749	245,680	10.4%
7	Canada	262,627	189,924	183,093	223,214	-5.3%
8	Australia	227,995	170,035	185,870	185,663	-6.6%
9	Hong Kong	148,644	151,066	152,294	161,018	2.7%
10	Philippines	63,412	72,486	80,838	131,788	27.6%
11	Mexico	123,368	109,405	95,473	123,495	0.0%
12	Brazil	124,587	77,826	68,016	105,982	-5.2%
13	Singapore	105,818	102,172	98,678	102,603	-1.0%
14	Malaysia	70,498	71,978	72,957	101,834	13.0%
15	Russian Federation	103,978	78,792	95,506	97,335	-2.2%
16	Israel	42,972	39,324	48,326	74,434	20.1%
17	Saudi Arabia	65,877	69,446	61,308	64,941	-0.5%
18	Colombia	73,628	64,090	57,778	55,906	-8.8%
19	Ukraine	60,051	53,051	50,516	42,514	-10.9%
20	Pakistan	19,776	22,970	23,395	38,250	24.6%
21	Switzerland	66,400	35,752	38,695	34,337	-19.7%
22	Egypt	71,705	63,989	45,794	31,390	-24.1%
23	India	16,377	19,843	20,333	21,076	8.8%
24	New Zealand	21,533	21,735	21,087	17,951	-5.9%
25	Chile	12,497	10,848	14,090	17,506	11.9%
26	UAE	36,654	10,414	14,011	16,381	-23.5%
27	South Africa	3,096	2,599	8,888	15,423	70.8%
28	Cambodia	14,524	16,094	12,384	13,806	-1.7%
29	Dominican Republic	11,573	14,182	12,242	12,462	2.5%

Source: own processing, ITC

Table 5: General information of fishery exports from Vietnam and the 29 selected importers.

Variable	Unit	Obs.	Mean	Std. Dev.	Min	Max
EXPORT	US$ Thousand	116	206,495.80	350,691.70	2,599.00	1,710,337.00
GDP_EXPORTER	US$ Billion	116	202.13	14.30	186.20	223.78
GDP_IMPORTER	US$ Billion	116	1,801.58	3,808.14	16.70	19,390.60
COST_EXPORT	US$	116	443.25	8.26	429.00	448.00
COST_IMPORT	US$	116	552.00	308.00	0.00	1,554.00
DISTANCE	Days	116	18.52	11.38	2.00	43.00
OPENNESS_EX	Percentage	116	183.34	11.27	169.53	200.38
OPENNESS_IM	Percentage	116	91.78	84.04	24.12	425.98
FDI	US$ Million	116	692.09	1,695.92	0.00	8,937.78
TARIFFS	Percentage	116	5.64	6.95	0.00	30.00
FTA	Dummy	116	0.41	0.49	0	1

Source: own processing

Table 6: Descriptive summary of data.

	---- Coefficients ----			
	(b	B)	(b-B)	sqrt(diag(V_b-V_B))
	fixed	random	Difference	S.E.
LNGDP_EXPORT	2.4780	3.3695	-0.8915	0.6177
LNGDP_IMPORT	0.9562	0.6635	0.2927	0.3755
LNCOST_EXPORT	-3.4659	-3.7203	0.2544	2.1250
LNCOST_IMPORT	1.0320	-0.7206	1.7526	1.1894
LNOPENNESS_EX	-3.4426	-4.4891	1.0464	1.1585
LNOPENNESS_IM	-0.1426	-0.1502	0.0076	0.5057
LNFDI	0.0394	0.0370	0.0025	0.0051
TARIFFS	0.0028	-0.0028	0.0056	0.0037

b = consistent under Ho and Ha

B = inconsistent under Ha, efficient under Ho

Test: Ho: difference in coefficients not systematic

$chi^2(8)$	=	(b-B)'[(V_b-V_B)^(-1)](b-B)
	=	7.92
Prob>chi^2	=	0.4418

Source: own processing

Table 7: Results of the Hausman test.

Breusch and Pagan Lagrangian multiplier test for random effects

" LNIMPORT[pairid,t] = Xb + u[pairid] + e[pairid,t]"

Estimated results:

	Var	sd = sqrt(Var)
LNIMPORT	1.7032	1.3051
e	0.0395	0.1988
u	0.5530	0.7437

Test Var(u) = 0

chibar2(01)	=	86.17
Prob > chibar2	=	0.0000

Source: own processing

Table 8: Result of Breusch and Pagan Lagrangian multiplier test for random effects.

Volume XII Number 4, 2020

Applying GIS Technologies for Mapping Natural and Anthropogenic Transformed Soils in the Southern Forest-Steppe of the Republic of Bashkortostan

Marat Ishbulatov, Ilnur Miftakhov, Zila Churagulova, Radik Mindibayev, Alexander Komissarov

Federal State Budgetary Educational Establishment of Higher Education "Bashkir State Agrarian University", Ufa, Russia

Abstract

Soil is a core element of the biosphere, and the soil cover is exposed to major processes that take place within this system. Therefore, it is very important to understand the results of soil research from the perspective of current global and local environmental problems. With the advent of new methods of spatial analysis and techniques for obtaining remote sensing data, geoinformation technologies offer great opportunities for analyzing the natural and ecological state of the region. Therefore, geoinformation analysis of the state of natural and anthropogenic transformed soils is an essential tool for their studying, forecasting the development of the natural environment, and working out the ways of rational farming. As forest territories have a beneficial effect on the natural and climatic situation in general, and vice versa, environmental degradation on agricultural lands will result in worsening the situation on forest lands, it is necessary to consider forest and agricultural areas in close integration with GIS technologies. It should be done to improve the overall natural environmental conditions. The study used soil survey data conducted in 2017-2019. Field and office studies were conducted: samples were taken in the field to determine the agrochemical parameters of the soil, and the data obtained were analyzed using mathematical and statistical methods. Digital cartographic materials were created using geoinformation technologies. The basis for a comprehensive natural and environmental assessment of forest and agricultural areas using geoinformation systems was laid. The studies conducted to identify changes in natural and anthropogenic transformed soils have shown that the contours of soil varieties have changed. In many cases, there is a deterioration in soil properties. The number of fertile chernozem has decreased. In areas with low crop cultivation, there are signs of a decrease in the humus horizon and the development of erosion processes. As a result of the conducted research, a single digital soil and geographical database for forest and agricultural territories were created. The developed methodology and algorithm for creating a database and digital cartographic basis using geoinformation technologies in environmental studies can be recommended as a base for similar studies both in the Republic of Bashkortostan and in other regions

Keywords

Environmental conditions, forest and agricultural territories, GIS technologies, soil mapping.

Ishbulatov, M., Miftakhov, I., Churagulova, Z., Mindibayev, R. and Komissarov, A. (2020) "Applying GIS Technologies for Mapping Natural and Anthropogenic Transformed Soils in the Southern Forest-Steppe of the Republic of Bashkortostan", *AGRIS on-line Papers in Economics and Informatics*, Vol. 12, No. 4, pp. 29-45. ISSN 1804-1930. DOI 10.7160/aol.2020.120403.

Introduction

The leading role is assigned to the effective use of land resources, regular and systematic improvement of soil fertility to solve the food problem successfully. The development of modern geoinformation technologies allows analyzing the soil state and visualizing it in the form of the map that reflects each assessment stage, as well as the synthesis of the results obtained. The end of the 20th century and the beginning of the 21st century is an unprecedented era of low food prices for Western industrialized countries. The main challenges facing humanity today are overpopulation, food shortage, environmental pollution, deforestation, and natural disasters. According to the reports of the UN Food and Agriculture Organization, deforestation continues worldwide at a rapid pace. Every year, 15 mln ha of forest is lost, while only 5 ha grows.

According to scientists, the planet overpopulation can lead to global conflicts due to food shortages (FAO, 2011). Currently, the desertification processes in the South of Russia affect about 60 mln ha of farmland. One of the most significant desertification factors in this area is agricultural land degradation and the degradation of natural pastures in particular. In the South of the European part, the chernozem soils of Kalmykia and the flat areas of Dagestan have been particularly affected. Currently, these regions are characterized by the most intense manifestation of anthropogenic desertification, which covers more than 70% of the territory. In Astrakhan, Volgograd, Saratov, Samara regions, and Tatarstan, up to 50% of the entire territory is affected by desertification (Pendzhiev, 2013). The geographic decision support system is a very complex area since a considerable amount of spatial data is accumulated in various fields of application, from remote sensing to GIS, computer mapping, environmental assessment, and planning. GIS will provide an opportunity to create maps, integrate information, visualize scenarios, solve complex problems, introduce powerful ideas, and develop practical solutions. Using the GIS application, you can open digital maps on your computer, create new spatial information to add to the map, create printed maps customized to their needs, and perform the spatial analysis. Spatial database management systems aim to make spatial data management easier and more natural for preparing maps for users or applications such as urban planning, utilities, transportation, and remote sensing (Manjula et al., 2011).

Soil is a natural body consisting of layers (soil horizons) resulting from the interaction of climate, topography, organisms, source material (underlying geological rock), and time (Lucà et al., 2018). Accurate soil data are needed to develop reliable, high-resolution soil maps for the hydrological analysis, environmental protection, agriculture, and forest management. The information elements required for the soil characterization obtained from the sources with different spatial resolutions can be easily stored and managed within the framework of a geographic information system (GIS). Digital soil mapping (DSM) allows analyzing the relationship between soil properties and auxiliary data (for example, cartometric attributes and remote/proximal sensing data) using several pedometric techniques. The research is an attempt to use GIS in soil science at various spatial scales by describing the growing availability of auxiliary data for soil characterization; illustrating the primary relationships between soil properties and digital terrain models – derived topographic objects; generalizing spatial and non-spatial pedometric techniques for the analysis and modeling of soil properties; and illustrating some GIS applications in soil science, as well as using the obtained data in land use planning and soil protection.

Due to the wide variety of soil types and the complexity of soil and environmental conditions in the southern forest-steppe of the Republic of Bashkortostan, specific zonal and locally differentiated and scientifically based approaches to maintaining and improving the soil fertility are required. These conditions are also imposed by economic activity, which has a significant impact on the vegetation and soil cover. Deforestation, plowing vast land areas, unregulated livestock grazing leads to the destruction, reduction, and thinning of vegetation cover, all this has a negative impact on the soil condition of the region. For example, according to the statistical reporting, in 2019, the results of satellite monitoring of the forests in the Republic of Bashkortostan revealed 20 violations of illegal deforestation ith a total volume of about 3,500 m3 of wood.

The deterioration of the general environmental conditions due to the energy and mass exchange imbalance in the "soil-plant-environment" system and the observed weakening of the ecological functions of soil fertility, decrease in the accumulative horizon thickness, negative humus balance, genetically low nitrogen and phosphorus content, destruction, and preconsolidation of the arable soil layer, accompanied by a marked decrease in the air and heat transfer and biological activity of soils are characteristic features of the current state of agricultural land (Zaripova et al., 2009).

Agricultural production is a key force affecting soil processes and functions. Due to biophysical constraints, as well as rapid structural and technological development, new methods of agricultural production management appear that have a significant negative impact on soil processes and functions (Techen et al., 2020). This makes it difficult to assess the potential of land resources to meet the growing demand for food and non-food products. From the point of view of rational soil use, sustainable intensification means that the volume of products received from the given territory increases while maintaining soil quality indicators such as humus horizon, humus and nutrient content, environment pH, and salinity. The research aims to study the impact of new land management methods on soil processes. There are four categories of land management methods: scientifically-based farming system, crop rotation design based on soil and topography, use of reasonable mechanisms for soil treatment, and soil pollution. At the same time, they study the efficiency of nutrient use compared to traditional crop cultivation systems, the state of the soil-rhizosphere microbiome and its interaction with crops in crop rotation, the effect of soil compaction on soil-plant-atmosphere interaction, and the ecotoxicity of plastics, pharmaceuticals, and other pollutants that enter the soil. There is a need for an interdisciplinary, systematic approach to soil science and research related to process modeling, data management, and assessment of soil stability under various management methods. The identification of soil science research problems from the point of view of agricultural management contributes to the establishment of cooperation between various scientific disciplines in the field of sustainable agricultural production.

Our planet is experiencing one of the fastest climate changes in the history of the Earth. The current change is particularly significant, as it is most likely a consequence of human activity since the 19th century (Jia et al., 2019). The digital Earth model, which includes satellites in near-Earth orbit, ground-based observations, and other data collection, analysis, and visualization technologies, has allowed scientists to see our climate and its impact at the regional and global levels. The Digital Earth platform provides valuable information about the atmosphere, biosphere, hydrosphere, and cryosphere to understand the past and present of the Earth. It also supports Earth system models for climate prediction and forecasting. Case studies where the Digital Earth model is used in the climate change research, such as climate sensing, information, and simulation systems for global environmental change, and synchronous satellite, aerial, and ground observation experiments that provide extensive and abundant data sets. Mapping climate extremes and impacts increase the preparedness for climate change risks and provide reliable data for management decisions. However, Digital Earth faces challenges in coordinating and integrating multi-source data, which requires an international partnership between governments and other intergovernmental organizations to develop open data policies and practices.

With climate change, the problems of agricultural production and food security have become particularly acute in the twenty-first century. For example, the Himalayan country of Bhutan is an agricultural country where about 57% of the population depends on agriculture. However, farming is constrained by mountainous terrain and rapid changes in environmental variability. The country is already experiencing some of the effects of climate change, such as crop loss due to unusual outbreaks of diseases and pests, erratic rains, hurricanes, hail storms, droughts, flash floods and landslides annually (Chhogyel and Kumar, 2018).

One of the most serious problems in sub-Saharan Africa is the need to increase crop production to meet the increasing demand of the growing population. To this end, knowledge of soil resources and their agricultural potential is essential for determining proper and appropriate land use and land management (Nguemzei et al., 2020). Studies of soil fertility in the Tombel region were conducted to study the current state of soil resources and monitor the impact of physical and chemical properties on soil fertility. Analyses on various indicators revealed a direct influence of the physical and chemical soil properties and other derived parameters of soil fertility on the main factors of plant growth and development, such as the ability to retain water, root development, soil aeration, nutrient existence and availability, and cation balance. Based on the physical and chemical soil properties, fertility parameters, and the soil quality index, four classes of soil fertility were determined in this area: 1) very good fertile soil (66 km^2), 2) good fertile soil (506 km^2), 3) fairly good fertile soil (787 km^2), 4) low-fertility soil (375 km^2). The main indicators that control the soil quality in the Tombel area, obtained on the basis of ANOVA and PCA analysis, are Ca, Mg, water pH, organic matter, total nitrogen, and zinc. Four out of seven indicators (Ca, pH, OM, P) were also identified as important indicators for assessing the fertility state of various soil groups in the Tombel area.

Economic activities, including logging, coal mining and other types of work, are the most important disturbing activities that affect the ecological functioning and conservation of forest biodiversity, and also adversely affect the agriculture (Kimaro and Lulandala, 2013). In addition, it is noted that the values of species diversity, composition, and regeneration potential within undisturbed forest areas differ significantly from those in severely disturbed areas. These observations confirm that current human activity has already led to a deterioration in the quantity and quality of useful plants, an increased impact of species diversification on the forest ecosystem, and possibly a negative impact on the livelihoods of local communities.

Human activity leads to changes in the global environment, sometimes with serious consequences for our future lives (Várallyay, 2010). Changes in the gas composition of the atmosphere, partly due to the emission of CO2 and "greenhouse gases", can lead to an increase in temperature with high spatial and temporal variability, to changes in global circulation processes, and a serious precipitation redistribution, an increase in aridity in some areas. These changes have a sensitive effect on ecosystems (natural vegetation and land use structure) and significant changes in soil formation and degradation processes, as well as in soil properties and functions.

All indicators of the forest influence on climate, hydrology, soil formation, sanitary and hygienic properties, and recreational suitability of its growing area can be safely attributed to the ecological potential. It is established that the landscape method is subjective, it does not give an idea of the recreational suitability, does not allow for mathematical data processing, but can serve as a basis for preliminary object characteristics (Sultanovaet al., 2018).

The ecological state in India is significantly deteriorating, and soils are formed through the diverse interaction of several forces, including climate, terrain, parent species, and organisms. It takes thousands of years for soil to form, and most soils continue to transform after changes in some soil-forming factors, in particular, climate and vegetation over the past few decades. Climate is one of the most important factors affecting soil formation, which is essential for their development, use, and management in terms of soil structure, stability, moisture retention of the topsoil, nutrient availability, and erosion. Various authors predict that the expected changes in temperature, precipitation, and evaporation will cause significant changes in both organic matter turnover and CO^2 dynamics (Karmakar et al., 2016; Uskov and Bulat, 2014).

Soil erosion is a serious problem that humanity faces today, as it continually worsens the quality and standard of living around the world. As a hilly country with undulating terrain with steeper slopes accompanied by heavy rainfall, Nepal is prone to natural disasters, including soil erosion (Chalise et al., 2018). In the work of Chalise and colleagues, an attempt was made to model the soil erosion rate in the Aringale Khola catchment area of the middle hills of Nepal using the methods of the geoinformation system (GIS) and the revised universal soil loss equation (RUSLE). RUSLE was used in ArcGIS, taking precipitation – Runoff (R), soil erodibility (K), topography (LS), crop management (C), and other factors as primary input data. In total, nine classes of soil erosion were observed, the soil erosion rate ranged from 0.03 to 100.33 t/ha/year, with an average level of soil erosion of 11.17 t/ha/year. GIS analysis showed that 36.93% (1256.28 ha) of the total catchment area is highly eroded, while the remaining 63.07% (2145.56 ha) is less eroded. Much of the catchment area, especially areas with higher elevations and steep slopes, is degraded and needs urgent soil conservation measures. This study is the first attempt to model the spread of soil erosion in the Aringale Khola catchment and can be used to predict soil erosion in similar catchments in the middle hills of Nepal.

Soil cover is an important element in the study of the relationship between human activity and the environment. The monitoring of soil indicators is necessary to identify deviations to preserve the maintained environment (Kayet and Pathak, 2015). Remote sensing is a tool for monitoring land use and land cover. The studies were conducted to assess the state of land use and changes in the soil and vegetation cover at the beginning of 1992, 2005, and 2014 on the territory of the Saranda forest. ArcMap GIS software was used to create and analyze thematic maps. This software made it possible to identify land-use features, taking into account the soil and vegetation cover, to preserve a stable environment while increasing productivity. The study revealed a rapid expansion of the built-up (mining) territory and an increase in agricultural land area with a decrease in the area of dense forests and water bodies.

Deforestation and the conversion of natural pasture land to agricultural land are two major threats to soil and water conservation, causing erosion and possibly desertification. The aim of the research in the Tzicatlacoyan area was to assess soil erosion using the universal equation of soil loss through geographical information systems (GIS) (López-García et al., 2019). The results showed that Tzicatlacoyan faces the risk of soil erosion at an average annual rate of 117.18 t/ha/year, due to natural factors and anthropogenic activities, such as the use of agricultural land without erosion control measures. Four classes of soil erosion risk were identified depending on the erosion rate (a) in t/ha/year: extreme risk (114 ≥ a ≤ 234.36), severe risk (59 ≥ A <114), moderate risk (23 ≥ a <59), and low risk. Most of the area (180.96 km^2, 64.83%) was characterized by a low risk of erosion, while a small part (11.64 km^2, 4.17%) of the study area was characterized by a high risk. The results showed that on 13.33% of the Tzicatlacoyan territory, the current values of soil losses exceed the permissible ones. Soil erosion assessment using a GIS model can allow land users to make more effective decisions about land use while preserving the soil and the entire ecosystem.

To identify changes in soil salinity and their impact on the vegetation cover is necessary to understand the relationships between these changes in the vegetation cover. This study, aimed at determining the changes in soil salinity and vegetation cover over the past 28 years, was conducted in the Al-Ahsa Oasis. Landsat time-series data for 1985, 2000, and 2013 were used to obtain images with vegetation index (NDVI) and soil salinity index (SI), which were then used in image differentiation to identify the vegetation and salinity changes-no changes over two periods. Soil salinization in 2000-2013 was significantly higher than in 1985-2000, and the vegetation cover decreased to 6.31% over the same period (Allbed et al., 2018).

In Turkey, the information about the current environment as a result of anthropogenic impacts on nature is recorded in an electronic environment, and an anthropogenic biome map has been created using the ArcGIS Desktop software. Analyzing the data obtained, we can say that over the past two centuries, the natural habitat has undergone considerable changes, the vegetation has deteriorated, and land degradation has increased due to the anthropogenic activity (Curebal et al., 2015).

The cadastre reflects the land plot boundaries of various owners. Changes in land legislation in recent years require considering and evaluating land by use type and soil varieties (Shapovalov et al., 2018). This method differs significantly from land accounting and valuation by land category and owner. As a result, the dynamics of cadastral division occurs in the form of an increase in the number of cadastral units – land plots, which is reflected in the public cadastral map. On the other hand, there is no real monitoring of land use, which leads to significant discrepancies with cadastral data and real changes in land use, which are detected during the retrospective monitoring of soils and vegetation cover. Real changes in land use structure are better correlated with soil cover than with cadastral division. The conflict resolution between the soil map compilation, cadastral division, and real land use can be achieved by introducing the soil and land cover concept. It is assumed that land cover maps will be created using retrospective monitoring technology.

Soil mapping and evaluation work allow compiling a graphical and descriptive database that is necessary for the inventory, classification, and evaluation of soil resources, which can be represented by an agricultural enterprise or administrative territory. The topsoil is studied in connection with natural and anthropogenic factors that determine its features and, accordingly, natural fertility, with various signs of favorability for the growth and development of agricultural or natural phytocenoses (Gosa and Mateoc-Sirb, 2014).

The use of geo-information systems (GIS) and remote sensing (RS) methods in soil mapping and classification provides significant advantages both in terms of time and cost and in terms of improving accuracy (Ramazanoglu et al., 2019). The study of Turkish scientists aimed to determine the conditions of soil formation, as well as to compile a classification of soils and their properties using GIS methods for the catchment of the Topchu stream in the Tarsus state, Mersin. Following the detailed soil geodetic and cartographic standards, a series of soils were outlined, soil samples were taken along the horizon, the physical and chemical soil characteristics were analyzed, and their classification was carried out. The change in topography, source material, land use, and surface properties of soils was determined using geoinformation methods.

In connection with the above, it should be noted that it is necessary to consider the soil and environmental situation in a particular region together on the territories of the forest fund and agricultural lands since the reduction of both forest and vegetation cover adversely affects the environmental situation as a whole.

The purpose of our study is to analyze the state of natural and anthropogenic transformed soils on the territory occupied by forests and crops, using geoinformation technologies for natural and environmental assessment and ensuring the sustainable development of natural and anthropogenic systems.

In this regard, the following tasks were set:

1. Create a single digital soil database for agricultural and state-owned land.
2. Develop a methodology for creating soil maps and their content using modern geoinformation technologies.
3. Conduct a comparative analysis of soil conditions for agricultural and state-owned land.

Materials and methods

The objects of the research were natural and anthropogenic transformed soils of the Dyurtyulinsky district of the Republic of Bashkortostan. The research program included the following: preparatory work, field soil mapping, laboratory analyses, office data processing.

Preparatory work consists of various, interrelated actions designed to ensure the quality and high efficiency of the entire study. During the preparatory period, the following materials have been selected and studied:

- aerial photography materials;
- previous soil survey materials;
- reports of soil and climate zoning of the survey territory;
- land assessment materials (soil and economic assessment);
- as well as literature and reference materials.

As a result of the analysis during the preparatory work, the primary materials of the soil survey and aerial photography data, when applied to MapInfo software, gave the result for compiling a preliminary layout of the updated soil map.

In field studies, a preliminary model of the soil map and aerial photographs were used as the planning and cartographic basis, aerial photographs are used to clarify the boundaries of individual soils and complexes. Fieldwork was carried out using the methods adopted in soil science, set out in the relevant guidelines (Ishbulatov et al., 2018a). Field studies were conducted on a scale of 1: 25000. The main sections, half-pits, and heeling-in were laid at the rate of one soil section per 5 hectares of the area. Soil samples at each point were selected based on the genetic horizons. Soil analyses were carried out according to generally accepted methods and Russian National Standards.

Research design

Analytical data processing was carried out using generally accepted variational and statistical methods based on the guidelines (Kalinin, 2015). The cartographic research method was previously widely used in the thematic map compilation. It included work on the collection and systematization of stock materials, as well as field expedition, laboratory, and office work. In the office period, cartographic work was performed using the terrain plasticity method, which allowed displaying the system and structural organization of delta surfaces on the map. Currently, this method has been replaced by geoinformation technologies that allow you to display all map layers on a single screen, and also accelerated the soil mapping process. The geographic information system (GIS) is a software and hardware complex based on digital maps with databases linked to them. GIS consists of two large blocks: electronic maps with databases and tools for ensuring GIS functioning. The use of GIS allows us to transfer the problem solution on creating a land assessment basis to a new qualitative basis (Ishbulatov et al., 2018b).

The essential GIS advantages:

- ability to automate the process of creating maps;
- ease of making changes, the ability to create systems for automatic adaptions to the database.

Equipment

The material was processed on a PC using standard computer programs such as "STATGRAPHICS Plus", "MS EXCEL", Mapinfo, Qgis.

Agroclimatic

Diurtiulinsky district refers to the Right-bank and partly to the Left-bank Pribelskiy hillside-plain district. It is characterized by the development of rocky and hilly plains with vast watersheds. The territory located on the watershed of the Belaya and Tanyp rivers, as well as in the valley of the Belaya River, has a high plowing rate (62.96%) of the territory.

Diurtiulinsky district is located in the lower reaches of the Belaya River. There are two specially protected natural areas (SPNT) on its territory: the medicinal plant population in Dyurtyulinsky and Kangysh forestries. They, along with other SPNT, are recorded in the Register of specially protected territories of the Republic of Bashkortostan. There are forest-seed reserves and plantations of unique pinetum (Churagulova, 2003).

The study area of the Republic of Bashkortostan is located in a zone of warm-temperate semi-arid climate. The zone is characterized by sufficient, but not always stable moisture. According to Verkhne-Yarkeevo and Dyurtyuli weather stations, the climatic conditions are very variable over the seasons. Sometimes there are abnormally hot days in summer, the highest temperature is +38˚C, for example, in 2010, there were abnormally cold days in winter, the lowest temperature was -46˚C, etc. The average annual air temperature here ranges from 2.3˚C to 2.5˚C, the average monthly temperature of January is -14.6˚C, of July from +18.3˚C to 19.1˚C. The amount of annual precipitation is from 350-400 to 450-500 mm. The sum of temperatures for the same period is 2000-2200˚C.

The beginning of the frost-free period varies over the years from mid-May to the first decade of June; the end is more often noted before the first half of September. The length of the frost-free period is from 100-110 days in the Northern part to 130 or more – in high places. The snow cover height in the first half of winter is 15-20cm, in the second one – 30-40 cm.

The area under study is characterized by a wide variety of vegetation. It includes many of the most diverse plants: large trees that raise their crowns many meters above the ground, relatively high shrubs, low shrubs, all kinds of herbs, including medicinal ones, and very small mosses and lichens. And each takes its specific place in the forest, plays a particular role in the forest life and life forms living in it. The forest is of exceptional importance not only as a source of wood production but mainly as a climate-forming factor that fulfills water conservation, soil, and wind protection. Forests that grow here are allocated to the area of dark coniferous-broad-leaved and pine forests.

Data analysis

As a result of the study, processing and analysis of the obtained qualitative and quantitative data were performed using GIS technologies. The results of the analysis can be presented as a map, values in a table, or chart-new information. It is necessary to decide what information should be mapped, how to group the values for the best data presentation.

In the process of result evaluation, the objectivity and necessity of the information obtained are determined to decide on repeating the analysis with other parameters or using another method. GIS makes it relatively easy to make the necessary changes and get a new result.

Results and discussion

From 1956 to 1997, on the instruction of the Ministry of Agriculture of the Bashkir Autonomous Soviet Socialist Republic, the employees of the Bashkir branch of the "Volgogiprozem" Institute conducted field soil surveys in all 54 districts of the Republic of Bashkortostan. From 2016 to the present day, Bashkir State Agrarian University, together with Volgoniigiprozem, are correcting the materials of the previous survey round, and electronic soil maps with layers in the geoinformation system are being created for the first time. At the moment, this work has been carried out in 33 districts. The purpose of the work was to study the soil state, identify changes that occurred after the last soil survey round, and compile a digital soil map. The soil survey is carried out in the municipal areas. If earlier these works were performed on the territory of individual agricultural enterprises, now the research is conducted in the context of rural settlements.

The following soil types and subtypes are identified on the territory of agricultural lands (117,229.17 ha) and the former Durtyulinsky specialized seed forestry (29,699 ha) (Table 1).

	Name of soil types and subtypes
1	Sod-weak-medium-strong podzolic, sod-podzolic gleyed
2	Light gray, gray, dark gray texture-differentiated
3	Clay-illuvial chernozem
4	Wet-meadow
5	Marshy, peaty-marshy
6	Gully-beam complexes
7	Floodplain alluvial humus, dark-humus, dark humus quasi-gley

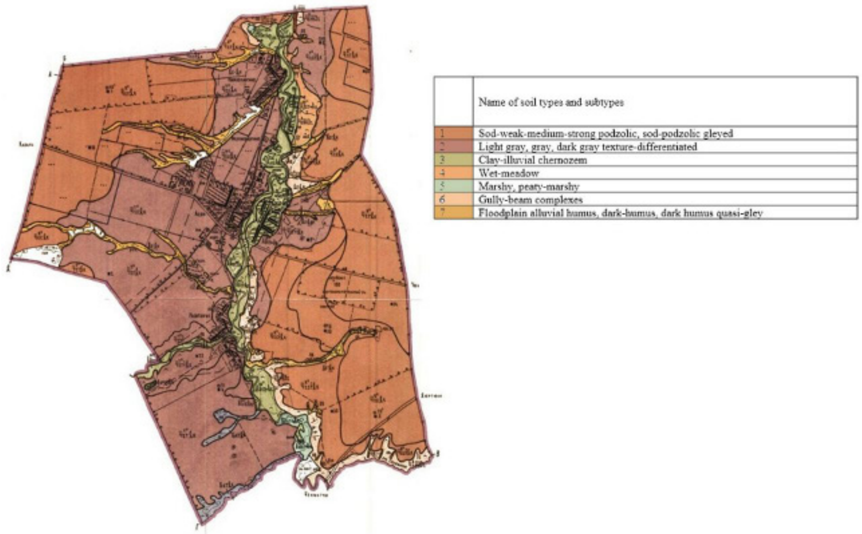
Source: authors

Table 1: Prevailing soil types in Durtyulinsky district.

The soil cover of the territory, as shown by the data of continuous mapping, is represented by sod-podzolic, light-gray, gray cinnamonic, gray, dark-gray forest, black soil podzolized, meadow-chernozemic, meadow-bog peat, alluvial floodplain granular, dark-grained sod-gleyed, and other soils, on the areas subject to anthropogenic impact – agro-soils with a homogeneous horizon.

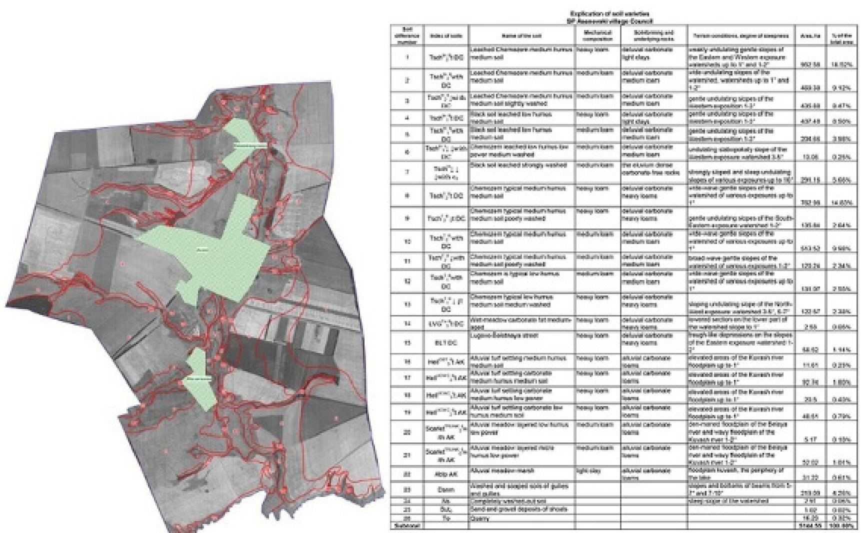
The use of GIS technologies in soil mapping is primarily related to the digitization of map material. Several digitization techniques are used, depending on the available hardware and software products. The typical positions are scanning the topographic base and assigning coordinates to the resulting raster image, or decoding and digitizing orthophoto plans (halftones). The result is a map in Mapinfo format, which is converted to a raster base with reference to the local coordinate system (MSK 02 zone 1). There are two ways to digitize a paper original: manual contour outlining on a registered topographic basis, on a recorded, scanned map, or automatically using vectorizers that digitize a paper-scanned contour grid.

Based on the received field survey materials and laboratory research results, the soil maps are corrected and digitized. The soil contour digitization is carried out on the basis of soil maps made in the period from 1992 to 1994 (Figure 1 and 2).



Source: authors

Figure 1: Soil map made according to the results of the 1992 soil survey within the boundaries of the Asyan collective farm of Dyurtyulinsky district RB.



Source: authors

Figure 2: Orthophotoplan (halftone) flight 2007 for digitizing the planimetric base.

When digitizing, it is necessary to restore the information that is partially lost or distorted due to the paper media wear, as well as scanning errors. In this regard, when digitizing maps, it is necessary to use the author's copies of soil maps. Next, digital soil maps are corrected. The elimination of inconsistency and disalignment is carried out by means of office clarification using cartographic documents and soil survey data from previous years.

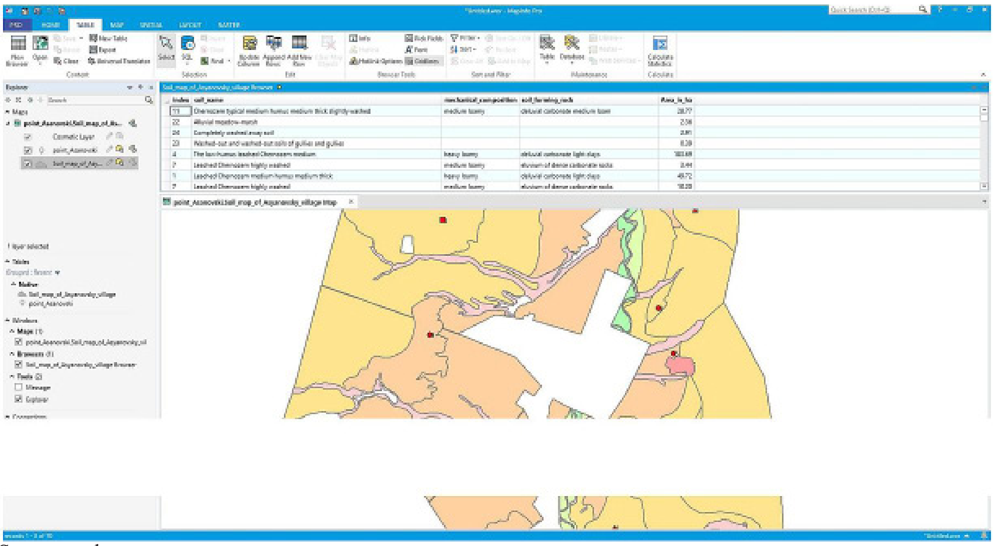
Simultaneously with the land map, the soil map obtained as a result of the adjustment is digitized.

Cartographic models for analyzing the soil and ecological state and spatial information are combined with an electronic database that contains multicomponent characteristics of the study area. This approach allows for land and environmental assessment to use its data in actions to take measures to improve the ecological situation. Creating a database in a GIS environment enables you to make adjustments about changes in various assessment subjects quickly. The soil database contains:

The results of this work are presented in the form of a complex electronic database.

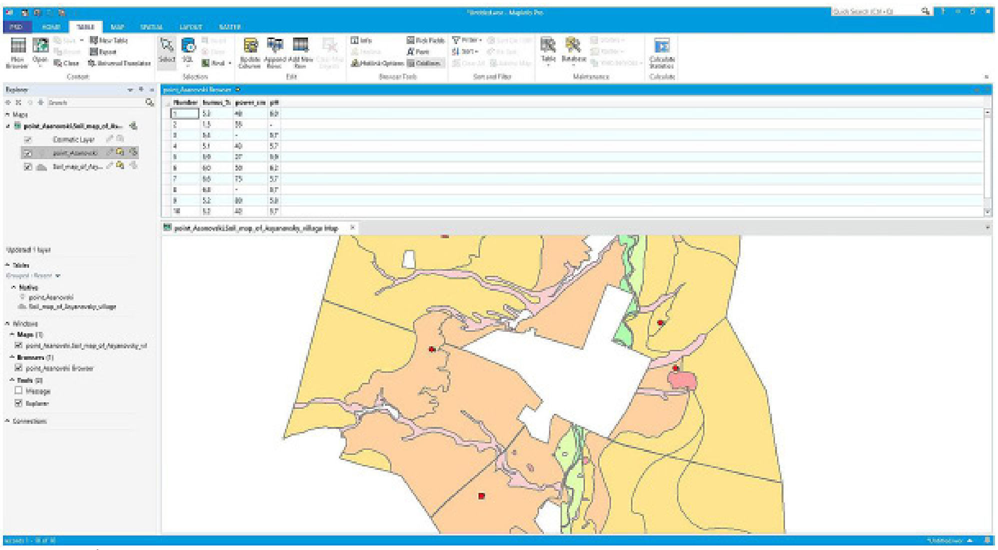
Layer 1 – soil index, soil names, granulometric composition, soil-forming, and underlying bedrock, and area (Figure 3).

Layer 2 – humus content in the soil, nutrient planetism, pH value, thickness of the humus horizon (Figure 4).



Source: authors

Figure 3: Electronic attribute database of the digital soil map, Layer 1.



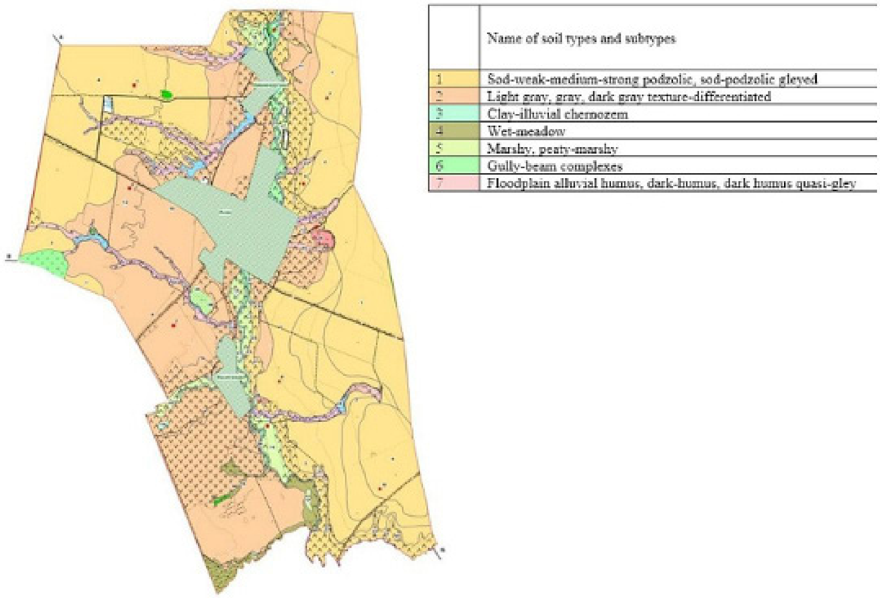
Source: authors

Figure 4: Electronic attribute database of the digital soil map, Layer 2.

A raster base that is put in correspondence with the local coordinate system (MSK 02 zone 1) serves as a planimetric base to open the created layers and get the soil information. An electronic soil map containing information about forest and agricultural lands is shown in Figure 5.

Field soil surveys conducted on agricultural lands and state forest fund showed that there are no significant differences in soil varieties.

The following offers the data for a comparative analysis of adjacent sectors (Tables 2 and 3).



Source: authors

Figure 5: Soil map, based on the 2018 soil survey, as a result of the digitization and adjustment of the soil maps within the boundaries of the Asyanovsky village council of Dyurtyulinsky district RB.

Soil phase number	Main section number, half-pits, and heeling-in	Sampling depth, cm	Thickness of the humus horizon, cm	Humus, %	Salt extract pH	Physical clay, %	Mobile phosphorus, mg / kg soil	Exchange potassium, mg / kg soil	Total absorbed bases, mmol / 100 g soil	hydrolytic acidity mg-eq / 100 g soil
1	1	5-15		4.3	7.4					
2	5	20-30	34	1.7	6.0		82.0	35.0	9.5	
		40-50		0.9						
		60-70		0.6						
4	3	5-15	18	1.0	5.5		75.0	50.0	9.9	2.0
		20-30		0.7						
		50-60		0.3						
6	4	10-20	38	2.4	5.9		60.0	60.0	30.5	1.5
		40-50		0.7						
9	2	10-20	33	1.4	5.9		47.5	50.0	38.7	
		45-55		0.7						
11	6	10-20	40	4.2	5.8		131.5	120.0	35.0	
		40-50		0.4						
12	9	20-30	29	2.7	6.5		37.5	60.0	52.7	
		40-50		2.2						
14	7	5-15	20	5.5	6.1		47.5	32.0	29.3	
		20-30		1.0						
	8	10-20		5.3	5.8					

Source: authors

Table 2: Indicators of agricultural land fertility status in Angasyakovsky village council.

Sample depth, cm	Humus %	Gross forms, %			Active forms, mg / 100g (according to Kirsanov)	
		N	P	K	P_2O_5	K_2O
Sandy loam. Section 1-86. Seedbed.						
Angasyak breeding nursery. Common pine (Pinus sylvestris), 2Y.						
0-26	2.7	0.25	0.12	0.3	4.6	14
26-36	1.6	0.125	0.075	0.88	8	16.5
45-55	0.6	0.07	0.15	0.3	4	6.6
Sandy loam. Section 6-86. Gramineous pinery.						
Angasyak forestry, sq. 41. Siberian spruce (Picea obovata), 3Y.						
2-12	4.1	0.780	0.11	0.4	5.2	8.0
12-25	0.5	0.125	0.10	0.4	8.65	6.3
26-36	0.4	0.125	0.065	0.5	12.0	11.0

Source: authors

Table 3: Humus content, gross and active forms of nitrogen, phosphorus, and potassium on forest lands.

Since the conditions of soil formation are identical, the main agrochemical indicators on different-purpose lands do not differ much. However, agricultural land is more susceptible to the anthropogenic effect. Approximately the same indicators show that agriculture in this area is done at the proper level without disrupting the technological process.

The information obtained from the 1960s and 1970s, which is commonly used, is no longer relevant to represent the reality of the current soil condition (Kumar and Geeta, 2009). Due to various changes in land use, crop management, intensive cultivation, combined with unbalanced fertilization, some fertile soils in the past have reached the status of degraded or unproductive land. Therefore, the current focus is on developing modeling approaches using new GIS and remote sensing methods as a possible option for reducing the cost factor. A holistic approach based on GIS has proved effective in converting useful subjective, qualitative, and categorical information into objective and quantitative one that serves modern requirements to update soil information.

The soil and agrochemical survey of the Shchelkovo training and experimental forestry nursery territory was carried out using geoinformation technologies (Martynenko et al., 2019). The purpose of this work was to conduct a soil survey of arable sod-podzolic medium – or light-loam soils of the forest nursery territory, as well as to assess the contamination of the production areas with weeds. The peculiarity of this study was collecting, systematizing, and processing the information focused on the use of modern geoinformation systems. The survey points were selected in agreement with the nursery administration and specified following the actual satellite images. The geographical reference of test points was carried out using global positioning technologies (GPS, GLONASS) based on the NextGIS mobile application (Android OS). The photo fixation of nursery fields was carried out using geotagging technologies, which allow implementing the received graphic information directly into the geographic information system (GIS). Soil samples collected in the field were analyzed in the soil laboratory to determine the main indicators of soil fertility. Based on the obtained data, the cartograms were developed using GIS, which provide the information about the original nutrient content of nursery fields, as well as about the most important indicators that characterize the soil-absorbing complex (soil acidity, degree of base saturation, etc.). The study of weed infestation allowed us to develop a thematic map that reflects the weed distribution in the nursery fields. All the collected information was combined into a comprehensive geo-information system designed on the basis of the quantum GIS shell. As a result, the developed GIS will contribute to the operational monitoring of soil fertility and provide information support for the agricultural equipment used in the nursery to grow planting material.

The availability of sufficient and accurate spatial data related to land resources is the basis for the sustainable development of agricultural production. The work was carried out to use remote sensing and geostatistical analysis tools within the geographic information system (GIS) to map the land potential and crop production for an area in the Qattara Depression, the Western desert of Egypt. The depression is located to the East of the Qattara Depression between 30° 10' 4" – 30° 20' 57" N latitude and 28° 32' 26" – 28° 52' 10" E longitude, covering 630 km^2 (63,000 ha) (Abbas et al., 2020). Thirty-seven soil profiles were dug up to 150 cm. Soil samples were collected from different horizons and analyzed for their main properties. The Applied System for Land Evaluation (ASLE) was used through the appropriate software to assess land potential and suitability. Landforms include sand slabs, dunes, depressions, sabkha (Arabic term for landforms formed when saline soils dry out in deserts and semi-deserts), and water bodies. 74.69 % of soils are "poor" (C4), 0.47 % are "good" (C2), 11.68 % are "satisfactory" (C3), 11.12 % are "very bad" (C5), and 0.47 % are "non-agricultural" (C6). The main limiting factors are the soil structure, its salinity, and alkalinity. Suitability classes (including 22 crops) can be divided into highly suitable (S1), suitable (S2), moderately suitable (S3), marginally suitable (S4), currently not suitable (N1), and potentially not suitable (N2) for 22 crops. The most recommended crops are date palm (Phoenix dactylifera) and tomatoes. The studied soils require precise management methods to be used for agriculture in the future (Abbas et al., 2020).

Digital soil mapping involves the creation and placement of spatial soil information obtained using field and laboratory observation methods in combination with spatial and non-spatial indicators. These maps represent the soil characteristics and parameters in digital form. The method for obtaining a digital soil map is computer-integrated, based on GIS software and agricultural knowledge. In order to compile a digital map of Herat soil classification, a cluster of various relief features (slope, slope exposure, micro-relief) obtained from the DEM (digital elevation model) of Herat province with a geological map of the specified territory was used in the study. As a result of the research, a digital soil map has been obtained showing different soils of Herat province, which are highlighted in different colours on the map, and areas with the same colours have similar soil properties. The work has made a significant contribution to solving the problem of the insufficiency of current quantitative and accurate soil data. Thus, the results obtained can be used in solving social, economic, and environmental issues.

The development of forest soil digital maps is a necessity both worldwide and at the national level. Currently, there are many maps of this kind and geographic information systems (Dincă et al., 2014). One of the examples is SIGSTAR 200, a system implemented by the Pedological Research Institute, ICPA Bucharest. Soil variants were grouped into a series of 32 soil types and subtypes by the Romanian soil classification system (SRTS 2003). Changes were introduced to the map using the pedological database of the Forest Research and Management Institute (ICAS) (2,665 soil profiles have been made over the past five years within the National forest inventory) and Rumanian geology and ecosystem GIS maps. Overlapping the boundaries of the forest areas and production units on the soil map allows the geographical soil distribution within the administrative boundaries of the forest fund. The information from this map, which includes various databases, can be used in forestry activities, in the preparation of different national or international projects, in the creation of other maps and databases, as well as in other economic or scientific activities.

A natural landscape map (of administrative districts) of the Republic of Sakha (Yakutia) was compiled on a scale of 1:2,500,000. GIS tools were used to link thematic cartographic documents to a topographic map containing contour lines, relief, and a hydrographic network (Kolejka, 2018). Data from Google Earth played a significant role in the subsequent mapping stages: 2D and 3D images were used to identify floodplains. The glacial landform units of the foothill territories and the main types of geological and geomorphological units were determined using both published data sources and Google Earth images. A map of natural vegetation has been compiled, taking into account the humidity regime and vertical zoning of the climate and soils. Digital cartographic layers in GIS were gradually combined into a synthetic natural landscape map, and the identified natural and landscape units were also included in the regional classification of Siberian landscapes. It is established that the hierarchical system of natural landscapes differentiates three classification levels: high (landscapes are distinguished by their geographical location into lowland and high-altitude landscapes), intermediate (landscape differentiation by the genesis and relief division in the corresponding climatic zone) and low (natural landscapes are determined by the vegetation covering the soil type that has developed on this geological substrate).

In Russia, it is long overdue to create a world-class soil attribute (profile) information database, which should become the basis for creating a system for monitoring the soil state and developing measures for their protection and rational land use. There is a need for a unified system for collecting and storing the information about soils, which would be open to the general use (Kolesnikova et al., 2010). Kolesnikova and co-authors have studied the concept and methodology of creating a soil-geographical database of Russia (SGDB) and developed the proposals for the structure and content of the soil attribute database, which is based on the concept of representative soil profiles. A list of classifiers has been compiled for data presentation formats in SGDB based on the existing soil morphology concepts, classifiers that characterize main physical and chemical soil properties.

The scientists from Volgograd State University study a landscape program that involves the formation of a land-use system aimed at connecting the protective forest belts with the geomorphological elements of the catchment area, relief, and dissimilarity of agricultural territories adapted to the dynamically balanced state of matter and energy within the landscape. This approach contributes to the development of a farming land assessment system through forest reclamation. This is due to the transformation (restructuring) of the qualitative and quantitative characteristics of energy and mass exchange. Consequently, radiation, thermal, soil, hydrophysical, and hydrodynamic processes also change. Thus, the area adjacent to the protective forest belt is a zone of the deterministic processes, while further away from the forest belt, the space is open for changing all characteristics. When assessing the land geoecology, the agroforestry landscape was considered as an agricultural landscape modification, which is formed and operated under the influence of the protective forest belts. The optimal organization of the irrigated agriculture should take into account the landscape heterogeneity of the territory. That was done through the interpretation of the space photos. According to the bioclimatic zonal indications, the agricultural landscape types of dry steppe and desert steppe were determined. Irrigated soils of the Volgograd region are located mainly in dry-steppe agro-forest landscapes on dark chestnut and chestnut soils within the natural reclamation areas of the Volga and Ergeninsky hills and partially in the trans-Volga delta plain; in semi-desert agro-forest landscapes on light chestnut soils within the trans-Volga delta plain and the Sarpinsky lowland. The favorable hydrogeological and land reclamation situation on the territory of the Southern Volga upland makes it possible to revive the irrigation in the Volgograd region and thereby increase the productivity and sustainability of the agricultural production at a higher scientific level using a geo-ecological approach (Ruleva and Rulev, 2015).

Geospatial analysis of the selected territory (Western Turkey, Izmir region) with the combined use of Google Earth, Landsat TM satellite images, and Erdas Imagine GIS programs is presented in the article by Turkish scientists. Advantages of using satellite images (multispectral images of Landsat TM and ETM +, Google Earth) in combination with GIS software for geospatial analysis tasks that are often solved in Higher education courses in geography and Earth sciences. This work provides an example of a successful landscape study. Studying the distribution of various types of the earth's surface, modeled using Landsat TM and Google Earth, allows you to analyze the dynamics of landscape changes. The work has also demonstrated the effective use of the Google Earth web service for thematic mapping. Image processing technologies were used to classify the images using the methods of the available Erdas Imagine software modules. The Google Earth web service was used to check and validate the mapping results using the Google Earth binding module (Lemenkova, 2015).

However, the productive forces of soil resources are gradually being depleted on the planet. In many agricultural areas of Taiwan, the crop yields were destroyed as a result of severe soil erosion. Nitrogen fertilizers and fossil fuels are usually scarce, which leads to both increased agricultural production costs and increased geopolitical conflicts. The recent increase in demand for phosphorus has led to a sharp rise in the phosphate ore cost – from $ 80 per ton in 2000 to $ 450 per ton in 2015. Prices have fluctuated since then, and now it is about $ 700 per ton. In addition to rising costs, mining is also a difficult problem. According to rough data, Morocco has the world's largest geological phosphorus reserves, but most of them are located in the disputed areas.

On the other hand, there are only about 2% of the world's phosphate ore reserves. According to current production rates, the richest phosphorus sources in the United States will be depleted in 20 years, leading to the increased dependence on phosphorus imports, which will support the demand from agriculture and industry. Global warming, on the other hand, accelerates the release of microbial greenhouse gases (GHGs) in soils and plays a crucial role in ongoing climate change. Soils and forests are becoming the key factors for human survival and development. This global development trend should be a guideline for Taiwan's future land and forest development policy (Lu, 2017).

Machine learning methods are widely used to create digital soil maps. The map accuracy is partly determined by the measurement number and spatial location used to calibrate the machine learning model (Wadoux et al., 2019). However, determining the optimal sampling scheme for mapping using machine learning methods has not yet been considered in detail in digital soil mapping studies. The optimization of the sample design for soil mapping was investigated. The design is optimized using the spatial simulation annealing by minimizing the standard error of prediction (SEP). This approach has been applied in compiling the soil maps for Europe using subsamples of the Lucas dataset. Optimized subsamples are used as input data for the machine learning model, using a broad set of readily available ecological data.

Taking into account the above, the method of creating the digital soil maps for the entire territory, including those occupied by agricultural production and forest plantations, will allow for more efficient use of land resources.

Conclusion

The introduction of GIS technologies, along with the new methods of information collection, makes it possible to significantly simplify and automate the process of studying and monitoring land resources throughout the entire work chain. The digital information is easily editable and stores data without any time limits, being a matrix for obtaining the high-quality cartographic products of any content and design. The main problem is a certain complexity of the process of getting digital material, which is easily eliminated when improving the work methodology.

At this stage, the active GIS implementation in the process of managing land monitoring is constrained by several objective factors. First of all, this is a lack of digital material on land resources. GIS development in the Republic of Bashkortostan does not have a very long history, and a few organizations are engaged in maintaining digital spatial information at the national and regional levels.

At the moment, the primary task in creating a GIS is to accumulate the digital data bank, develop a unified methodology and requirements for obtaining digital information. Only after the material has been collected and consolidated over a sufficiently large area, all the features of geo-information systems can be used.

The work in this direction is necessary to provide users with further operational information, primarily rural producers. The world experience shows that information technologies are highly effective in the agricultural sector.

Corresponding authors
Marat Ishbulatov, Candidate of Agricultural Sciences
Department of Real Estate Cadastre and Geodesy, Federal State Budgetary Educational Establishment of Higher Education "Bashkir State Agrarian University", 50-letia Octyabrya Str., 34, Ufa, 450001, Russia
Phone: +79373107807, E-mail: ishbulatovma@rambler.ru

References

[1] Abbas, H., Abuzaid, A. and Yaccub, R. (2020) "Land evaluation of east Qattara Depression, Egypt using remote sensing and GIS", *5th International Conference on Biotechnology Applications in Agriculture (ICBAA)*, Benha University, 8-11 April 2020, Egypt.

[2] Allbed, A., Kumar, L. and Sinha, P. (2018) "Soil salinity and vegetation cover change detection from multi-temporal remotely sensed imagery in Al Hassa Oasis in Saudi Arabia", *Geocarto International,* Vol. 33, No. 8, pp. 830-846. E-ISSN 1752-0762, ISSN 1010-6049. DOI 10.1080/10106049.2017.1303090.

[3] Chalise, D., Kumar, L., Shriwastav, C.P. and Lamichhane, S. (2018) “Spatial assessment of soil erosion in a hilly watershedof Western Nepal”, *Environmental Earth Sciences*, Vol. 77. E-ISSN 1866-6299, ISSN 1866-6280. DOI 10.1007/s12665-018-7842-3.

[4] Chhogyel, N. and Kumar, L. (2018) “Climate change and potential impacts on agriculture in Bhutan: a discussion of pertinent issues”, *Agriculture & Food Security*, Vol. 7. ISSN 2048-7010. DOI 10.1186/s40066-018-0229-6.

[5] Churagulova, Z. S. (2003) “Soils of forest nurseries of the southern Urals: properties, changes, optimization”, TISSOT, Moscow.

[6] Curebal, I., Efe, R., Soykan, A. and Sonmez, S. (2015) “Impacts of anthropogenic factors on land degradation during the anthropocene in Turkey”, *Journal of Environmental Biology*, Vol. 36, No. 1, pp. 51-58. ISSN 02548704.

[7] Dincă, L., Sparchez, G. and Dincă, M. (2014) “Romanian's forest soils gis map and database and their ecological implications”, *Carpathian Journal of Earth and Environmental Sciences*, Vol. 9, pp. 133-142.E-ISSN 1844489X, ISSN 18424090.

[8] FAO (2011) “FAO in the 21st Century: Ensuring Food Security in a Changing World”, Rome, Food and Agriculture Organization of the United Nations.

[9] Gosa, C. I. and Mateoc-Sirb, N. (2014) “Agricultural lands bonitation and estimation of crop production in Almajului Valley, Caras-Severin County”, *14th International Multidisciplinary Scientific GeoConference SGEM 2014*, Vol. 2, Book 3, pp. 11-14. ISBN 978-619-7105-14-8. ISSN 1314-2704. DOI 10.5593/SGEM2014/B32/S13.002

[10] Ishbulatov, M. G., Miftakhov, I. R. and Mindibaev, R. A. (2018a) “Creating a digital soil map of agricultural lands in the Republic of Bashkortostan Col.: cadastral and ecological and landscape support of land management in modern conditions”, *Materials of the international scientific and practical conference of the faculty of land management and cadastre of VSAU*, pp. 100-103.

[11] Ishbulatov, M. G., Mindibaev, R. A., Safin, K. M., Baikov, A. G., Miftakhov, I. R., Baygildina, G., Zamanova, N., Khisamov, R. and Yagafarov, R. G. (2018b) “The Bioenergetic Approach to Evaluation of Arable Land Fertility”, *Journal of Engineering and Applied Sciences*, Vol. 13, pp. 8353-8359. ISSN 1816-949X.

[12] Jia, G., Zhang, L., Zhu, L., Xu, R., Liang, D., Xu, X. and Bao, T. (2019) “Digital Earth for Climate Change Research”, In *Manual of Digital Earth*, pp. 473-494. ISBN 978-981-329-915-3. DOI 10.1007/978-981-32-9915-3_14.

[13] Kalinin, A. G. (2015) “Data processing by mathematical statistics methods”, Monograph. TBI of SUCC, Chita.

[14] Karmakar, R., Das, I., Dutta, D. and Rakshit, A. (2016) “Potential Effects of Climate Change on Soil Properties: A Review”, *Science International*, Vol. 4, pp. 51-73. ISSN 1013-5316.

[15] Kayet, N. and Pathak, K. (2015) “Remote Sensing and GIS-Based Land use/Land cover Change Detection Mapping in Saranda Forest, Jharkhand, India”, *International Research Journal of Earth Sciences*, Vol. 3, pp. 2321-2527. ISSN 14373254.

[16] Kimaro, J. and Lulandala, L. (2013) “Human Influences on Tree Diversity and Composition of a Coastal Forest Ecosystem: The Case of Ngumburuni Forest Reserve, Rufiji, Tanzania”, *International Journal of Forestry Research*, Vol. 6, No. 2. E-ISSN 1687-9376, ISSN 1687-9368. DOI 10.1155/2013/305874.

[17] Kolejka, J. (2018) “Landscape Mapping Using GIS and Google Earth Data”, *Geography and Natural Resources*, Vol. 39, pp. 254-260. E-ISSN 1875-371X, ISSN 1875-3728. DOI 10.1134/S1875372818030095.

[18] Kolesnikova, V., Alyabina, I., Vorob’eva, L., Molchanov, E., Shoba, S. and Rozhkov, V. (2010) “Soil attribute database of Russia”, *Eurasian Soil Science*, Vol. 43, pp. 839-847. E-ISSN 1556-195X, ISSN 1064-2293. DOI 10.1134/S1064229310080016.

[19] Kumar, P. S. and Geetha, S. A. (2009) "A Holistic GIS-based Approach for Thematic Extraction of the Soil Erosion Estimates Using Predictors as a Function of Climate, Land Cover, Relief, Soil, and Topography", *Asian Journal of Water, Environment and Pollution*, Vol. 6, No. 3, pp. 73-78. E-ISSN 1875-8568, ISSN 0972-9860.

[20] Lemenkova, P. (2015) "Google Earth web service as a support for GIS mapping in geospatial research at universities", *Web-Technologies in the Educational Space. Problems, Approaches, Perspectives*, pp. 24. DOI 10.6084/m9.figshare.7211798. (In Russian).

[21] López-García, E., Torres-Trejo, E., López-Reyes, L., Flores-Domínguez, Á., Peña Moreno, R. and Olguín, J. F. (2019) "Estimation of soil erosion using USLE and GIS in the locality of Tzicatlacoyan, Puebla, México", *Soil and Water Research*, Vol. 15, No. 1, pp. 9-17. E-ISSN 1178-6221. DOI 10.17221/165/2018-SWR.

[22] Lu, S.-M. (2017) "Soil and Forest: The Key Factors for Human Survival", *Journal of Sustainable Development*, Vol. 10, pp. 105-119. I-SSN 1913-9071, ISSN 1913-9063. DOI 10.5539/jsd.v10n3p105.

[23] Lucà, F., Buttafuoco, G. and Terranova, O. (2018) "GIS and Soil", In *Comprehensive Geographic Information Systems*, pp. 37-50. ISBN 978-0-12-804793-4. DOI 10.1016/B978-0-12-409548-9.09634-2.

[24] Varma M. K. R., Jyothi, S., Varma, S. A. K. and Varma, S. V. (2011) "Construction of Spatial Dataset from Remote Sensing using GIS for Deforestation Study General Terms", *International Journal of Computer Applications*, Vol. 31, No. 10, pp. 26-32. ISSN 0975-8887.

[25] Martynenko, O., Karminov, V., Yugaj, E., Mutygullin, I. and Ontikov, P. (2019) "The use of GIS for agrochemical soil characteristics and weed infestation of Grebnevskiy nursery of Shchelkovskiy teaching and experimental forestry", *Dokuchaev Soil Bulletin*, Vol. 99, pp. 5-20. E-ISSN 2312-4202, ISSN 0136-1694. DOI 10.19047/0136-1694-2019-99-5-20. (in Russian).

[26] Nguemzei, C., Tematio, P., Yemefack, M., Tsozué, D. and Silatsa, Fs. (2020) "Soil quality and soil fertility status in major soil groups at the Tombel area, South-West Cameroon", *Heliyon*, Vol. 6, No. 2. E-ISSN 2405-8440. DOI 10.1016/j.heliyon.2020.e03432.

[27] Pendzhiev, A. M. (2013) "Ecological problems of desert development: migration, pasture improvement, and global land degradation", *Alternative Energy and Ecology*, Vol. 14, No. 136, pp. 89-107. ISSN 2348-5159.

[28] Ramazanoglu, E., Gundogan, R. and Saltali, K. (2019) "Soil formation, Classification, and Mapping of Soils of Catchment of TopçuCreek", *International Conference on Agriculture, Food, Veterinary and Pharmacy Sciences*, pp. 1122-1127.

[29] Ruleva, O. and Rulev, A. (2015) "Landscape Planning of Forest Amelioration on Irrigated Soils", *Geofrafia i geoinformatika,* Vol.11., pp. 74-80. ISSN 2519-4437. DOI 10.15688/jvolsu11.2015.2.9. (in Russian)

[30] Shapovalov, D. A., Rukhovich, D. I., Kulyanitsa, A. L., Kalinina, N. V., Koroleva, P. V. (2018) "Elements of soil and land cover mapping in the cadastre system of the Russian federation," *International Multidisciplinary Scientific GeoConference-SGEM*.

[31] Sultanova, R. R., Gabdrahimov, K. M., Khayretdinov, A. F., Konashova, S. I., Konovalov, V. F., Blonskaya, L. N., Sabirzyanov, I. G., Martynova, M. V., Isyanyulova, R. R. and Gabdelkhakov, A. K. (2018) "Evaluation of Ecological Potential of Forests", *Journal of Engineering and Applied Sciences*, Vol. 13, pp. 6590-6596. E-ISSN 1816949X, ISSN 18187803.

[32] Techen, A. K., Helming, K., Brüggemann, N., Veldkamp, E., Reinhold-Hurek, B., Lorenz, M., Bartke, S., Heinrich, U., Amelung, W., Augustin, K., Boy, J., Corre, M., Duttman, R., Gebbers, R., Gentsch, N., Grosch, R., Guggenberger, G., Kern, J., Kiese, R. and Vogel, H.-J. (2020) "Chapter four - Soil research challenges in response to emerging agricultural soil management practices", *Advances in Agronomy*, in press. Vol 164, pp. 179-240. DOI 10.1016/bs.agron.2020.01.002.

[33] Uskov, V. N. and Bulat, P. V. (2014) "Shock and detonation wave in terms of view of the theory of interaction gasdynamic discontinuities", *Life Science Journal*, Vol. 11, No. 8, pp. 307-310. ISSN 10978135.

[34] Várallyay, G. (2010) "The impact of climate change on soils and their water management", *Agronomy Research*, Vol. 8, No. 2, pp. 385-396. ISSN 1406894X.

[35] Wadoux, A., Brus, D. and Heuvelink, G. (2019) "Sampling design optimization for soil mapping with random forest", *Geoderma*, Vol. 355. ISSN 0016-7061. DOI 10.1016/j.geoderma.2019.113913.

[36] Zaripova, G. K. Kaipov, Y. Z., Safin, H. M., Mindibaev, R. A., Nugumanov, A. Kh. and Chanyshev, I. O. (2009) "Main technological methods for improving the soil fertility and effective use of degraded lands in Bashkortostan (recommendations)", Ufa, Print World, 44 p. ISBN 5961300757.

Volume XII Number 4, 2020

Using Data Envelopment Analysis in Credit Risk Evaluation of ICT Companies

Michaela Kavčáková, Kristína Kočišová

Faculty of Economics, Technical University of Košice, Slovak Republic

Abstract

The aim of the paper is to explore possibilities of diagnosis corporate credit risk through DEA and design an appropriate model for diagnosis of credit risk, which can be used in different sectors of national economy (e.g. agricultural, service sector or industry and innovation sector). The model differs from the conventional application of DEA because of variables selection and construction of production-possibility frontier. We illustrate application of models on sample 110 randomly selected companies during the 2013-2017 period. The reason for choosing the ICT companies is the fact that this sector is considered to be driving force behind the growth of the economy. The data has been obtained from Finstat. The results are divided into identification of 3 zones of corporate financial health with a different stage of credit risk. They show that DEA achieves a satisfactory value of a correct classification into the relevant zone (financial health, grey, and financial distress zone), but also the relatively high error rate of the DEA in the identification of companies in financial distress.

Keywords

Companies in the information and communication technology (ICT) services industry, credit risk, Data Envelopment Analysis, financial health.

JEL Classification: C14, G30

Kavčáková, M. and Kočišová, K. (2020) "Using Data Envelopment Analysis in Credit Risk Evaluation of ICT Companies", *AGRIS on-line Papers in Economics and Informatics*, Vol. 12, No. 4, pp. 47-60. ISSN 1804-1930. DOI 10.7160/aol.2020.120404.

Introduction

The companies in information and communication technology (ICT) services industry play an essential role in the economic environment as they have a significant impact on social and economic development. They change not only the interpersonal relationship but also the way the other industries do their business. As mentioned by Vanek et al. (2011), ICT development is driven by high dynamics that can even be surprising in many ways. The fast development of information and communication technologies lead to changes in job demand, roles, and general growth within the sector. According to Krausová (2018), the ICT sector is considered to be specific, as it is currently the driving force behind the development and growth of the economy, and at the same time represents a tool for transforming the traditional functioning of the society. That is why it is necessary to study the financial health of these companies and this way also the probability of bankruptcy of these companies.

Companies bankruptcy prediction is an ever-growing process of interest to investors or creditors, as well as borrowing institutions. As mentioned by Chaudhuri and Ghosh (2017) timely recognizing the companies´ moving toward failure is often desired. Bankruptcy can be considered as the state where the company is not able to satisfy the debts and requests of the court of low to restructure its debts or liquidate its assets. A timely prediction can help to evaluate risks in order to prevent bankruptcy.

Business bankruptcies are an essential part of everyday economic life. They happen every day around the world. At present, setting up their own business is not complicated and very costly, and therefore it is established in Slovakia only on average 18 500 per year. For this reason, it is crucial for the business itself to monitor its financial health and credit risk. Identifying and quantifying credit risk is essential in increasing the effectiveness, accuracy and consistency of risk management. It is a direct asset not only in credit approval but also in credit management, risk-based valuation, loan secularization and loan portfolio management.

Credit risk belongs to a group of fundamental financial risks that have a decisive impact on the success of various institutions. It is based on the uncertainty as to whether the counterparty will repay its (financial) liabilities promptly and at a fair amount, either due to inability or unwillingness to repay. It arises in the provision of loans, bank guarantees, project financing or investment transactions in financial derivatives. All transactions that are expected to be fulfilled at some stage (both financial and non-financial) from the external environment are subject to credit risk. This type of risk is the greatest risk for a bank in providing loans to other entities. For this reason, it is crucial for the bank to examine in detail the financial situation of the company, which plans to grant credit by analysing the financial health of the companies.

Evaluating the financial health, credit risk and determining the likelihood of bankruptcy is important for all sectors of the national economy, not just only for loans, but also for obtaining various state or EU funds and contracts. These funds and contracts are provided only to financially health companies that are not at risk of bankruptcy. A significant volume of grants is also intended for agriculture and forestry. Almost all support in Slovakia comes from the Rural Development Program, which invests in primary agricultural production, young farmers, construction of forest roads, purchase of forestry technology and much more. For this reason, it is very important for farmers and their strategic decision making and for policy makers to monitor their financial health. Various authors around the world address this issue within the agro-sector.

As mentioned by Kiaupaite-Grushniene (2016) the evaluation of financial health of agricultural enterprises is a important topic not only from an investor, but also from government and owner point of view. After accession to European Union, farmers are receiving significant subsidies from European and national budget, and those subsidies are intended to be provided only to financially healthy enterprises with future perspective of their sustainable economic performance. Therefore, there is a need to know how to distinguish te well performing companies from those who face the financial problems.

The question of financial healt of agricultural companies is also analysed by Zhao et al. (2008), who pointed to the fact that signalising of financial problems is an important element in the lender-borrower relationship that influences the cost and availability of debt capital, mostly from banks or from government, to agricultural borrowers. Therefore it is necessary to analyse the level of credit risk and the probability of bancrupty also in the companies from agricultural sector.

In Lithuania, Jedin and Stalgienė (2018) presented the likelihood of the Lithuanian family farms bankruptcy based on the economic size and type of farming by analysing financial indicators of farms. They used farm-level panel data for the year 2014-2016 from Farm Accountancy Data Network (FADN) and their research showed that more than 40% of small dairy farms had the high likelihood of bankruptcy in the year 2015-2016, as well as 30% of medium and large cereals, oilseeds and protein crops farms in 2016.

Specific credit scoring model for agricultural loan portfolio, which reflects major risk characteristics of Indian agricultural sector, loans and borrowers and designed to be consistent with Basel II, was developed by Bandyopadhyay (2007). In this study he shown how agricultural axposures are typically can be managed on a portfolio basis which will not only enable the bank to diversity the risk and optimize the profit on the business, but also will strenghted banker-borrower relationship and enables the bank to expand its reach to farmers because of transparency on loan decision making process.

In US Breer et al. (2012) analyzed the probability of default for USDA Agriculture Resource Management Survey (ARMS) farms over time. They used a synthetic credit rating model to predict the probability of default for each ARMS farm sampled. The farms are classified according to farm type, gross sales class and by region to assess the financial health of each sector. Results of these analyses provide insights into which farms may be under financial stress and whether those farms under stress have common characteristics.

Within the Czech Republic Karas, Režňáková and Pokorný (2017) dealt with the issue of predicting bankruptcy of agricultural companies, who analysed the current accuracy of four traditional bankruptcy prediction models (a revised Z-score model, Altman-Sabato's model – the version with unlogged predictors and the version with logged predictors, model IN05) in the field of agriculture. Their results showed that these models are less accurate in this field in comparison with the original results and state, that this motivates the effort of deriving new models that would be specially developed for agriculture business.

Currently, there are several models based on which a bank can determine whether a borrower is financially healthy or on the verge of bankruptcy. Financial health is a characteristic of the financial condition of those companies that show favourable proportions between the various aspects of their finances. A financially healthy company does not show any signs of jeopardizing its continued existence; from its condition, it can be assumed that it will not be insolvent in the near future. Indeed, such an undertaking shows sufficient profitability and adequate coverage of the risks of indebtedness. Based on several methods, such as Altman's Z-score, Kralick's Quick test, Beaver's model, IN indices, according to the achieved values of companies are classified into three zones, namely the zone of financial health, so-called grey zone and bankruptcy zone. These methods are called bankruptcy and creditworthy models.

According to Kotulič et al. (2007), models based on empirical-inductive indicator systems are used to predict the financial distress of an enterprise, using, in particular, the following methods:

- one-dimensional discriminatory analysis - mathematical-statistical method, which predicts the financial distress of a company using a single indicator (e.g., Beaver's model);
- multidimensional discriminatory analysis - also a mathematical-statistical method which predicts financial distress using a set of multiple indicators, with different weights assigned to these indicators (e.g., Altman's Z-score, IN indexes);
- scoring method - predicts the financial development of the company using scales, which are usually determined by expert methods (e.g. Kralick's Quick test).

One of the most frequently used traditional models is Altman's Z score (Altman, 1968), which belongs to the group of multidimensional bankruptcy models. This test is based on empirical data on failed companies over the past five years before bankruptcy and data for successful companies over the same period. Subsequently, he used Multiple Discriminatory Analysis to determine the ratios that characterize both the current financial situation of the companies and their development. The essence of this analysis is to find a linear combination of indicators that best distinguishes companies from bankruptcy and prosperity. One of the strengths of this model is that it allows for simultaneous consideration of several financial variables on the purpose of developing a bankruptcy prediction model. Also, it highlights factors contributing to a company's financial health and uncovers emerging trends that indicate improvements or deterioration in financial condition. On the other hand, the test uses unadjusted accounting data - it uses data from relatively small firms, and it uses data that is around 60 years old, or the test's predictive ability dropped off considerably from there with only 72% accuracy two years before failure, down to 48%, 29%, and 36% accuracy three, four, and five years before failure, respectively.

Beaver was one of the first authors, who analyse the company's financial health. His model is called the Beaver's model (Beaver, 1966), and it is based on the one-dimensional discriminatory analysis. The main disadvantage of this type of models is the fact that the same company can be included in the group of financially healthy, but also in the group of companies in financial distress based on different indicators. The authors try to solve this shortcoming by using the above-mentioned multidimensional discrimination analysis.

The last traditional model that we are going to mention belongs to the third group of traditional models used to assess the financial health of the company, namely Kralick's Quick Test (Kralicek, 1993). This model is often used mainly in Europe, and it is a kind of transition between one-dimensional and multidimensional models. From each significant area of analysis, such as stability, liquidity, profitability and economic outcomes, he used one selected indicator and created a point scale. The final score for the entire test is calculated as the average of the marks achieved for each indicator.

The main disadvantage of the all mentioned models is that the connection between the value of financial ratios of bankrupt and likely will change over time as obsolete and outdated. They need to be redesigned to the current economic situation, and each country should modify main models to its own conditions.

Many authors have applied these mentioned traditional models in their research. Some of the authors also modified them, such as Káčer et al. (2019), who investigated the classification performance of the re-estimated Altman's Z-score model for a large sample of private SMEs in Slovakia. They have found that even though the model with re-estimated coefficients achieves better discrimination performance, it is not statistically different from the revised Z-score model. Altman's Z-score was compared with other models, for example, Boďa and Úradníček (2019), who compared it with CH-Index and G-Index, Ékes a Koloszár (2014) compare it with logistical regression and neural networks or Vavřina et al. (2013) with logit models, DEA and production function-based economic performance evaluation. The approach of one-dimensional discriminatory analysis through Zmijewski's X-score have examined, for example, AlAli et al. (2018). The results obtained from this study have shown that companies working in that sector tend to have a healthy financial position, and they are on the safe side when it comes to bankruptcy risk. Husein, Pambekti (2014) or Aminian et al. (2016) have analysed the accuracy of the model of Altman's Z-score, Springate model, Zmijewski's X-score, and Grover model as the best predictor of financial distress. The results of first mentioned authors have shown that the model of Zmijewski's X-score is the most appropriate model to be used for predicting the financial distress because it has the highest level of significance compared to the other models. Aminian et al. (2016) concluded that the Grover model (Grover and Lavin, 2001) compared to the models of Altman's Z-score, Springate model and Zmijewski's X-score have shown better results. Since investors are always looking for knowledge the suitable situation for investment, and managers are interested in assessing the weaknesses and future threats and taking the necessary decisions in the face of these threats, they have suggested that the use of the Grover model as a tool for predicting bankruptcy or continuity of the companies and accordingly to make rational decisions.

However, these methods are outdated, and therefore scientists are trying to apply alternative methods that could better reflect the situation of the company in the current conditions. The multidimensional business performance environment is very attractive for the application of methods such as Data Envelopment Analysis (DEA), which takes into account both quantitative and qualitative information in the analysis. Data Envelopment Analysis is used to evaluate the technical efficiency of homogeneous units (Decision Making Units – DMU). It is a non-parametric method which belongs to a group of mathematical methods based on linear programming. This method aims to divide the surveyed objects into efficient and ineffective according to the size of inputs consumed and the number of outputs. DEA compares these objects to the best one.

The basis of DEA models is the so-called production–possibility frontier that is made up of all acceptable combinations of inputs and outputs. This frontier is determined by the production possibility curve, which determines whether or not the investigated unit is effective. The unit is effective if it lies at the production possibility curve. If it does not lie on this border, it is inefficient, and it is necessary to adjust the size of its inputs or outputs. By using DEA models, we can determine how to adjust these inputs or outputs in order to become an effective unit. Another basic term when using DEA is a Decision-Making Unit (DMU), which means the subject being evaluated in which the transformation process is taking place - the process of converting inputs to outputs.

DEA models can be classified according to different aspects:

- depending on how many inputs and outputs we are considering within the unit under review (evaluation of units with one input and one output, evaluation of units with two inputs and one output, evaluation of units with one input and two outputs, evaluation of multi-input and multi-output units);
- according to the orientation of the model (input-oriented, output-oriented, non-oriented);
- depending on whether they are based on the Constant Returns to Scale (CRS) assumption, the known Charnes-Cooper-Rhodes (CCR) model (1978), or the models based on the Variable Returns to Scale (VRS), e.g. Banker-Charnes-Cooper (BCC) Model (1984);
- radial (CCR model, BCC model and radial DEA models to calculate the so-called super-efficiency) and non-radial models (they monitor disproportionate changes in inputs and outputs for efficiency - ADD model of Charnes et al. (1985), Slack Based model (SBM) of Tone (2001)).

The most well-known scientist dealing with the application of DEA to various areas of the economy is prof. Paradi from the University of Toronto. This article will be based on his 2004 publication. In Slovakia, Mendelová and Bieliková (2017) used DEA to assess the financial health of industrial enterprises in Slovakia and compare the results with logistic regression and decision trees, inspired this publication.

Table 1 shows an overview of the studies applying DEA to evaluate corporate financial health or credit risk. Beside studies mentioned in Table 1, there are also many studies dealing with application of DEA in other sectors of national economy, for example in agricultural sector. For example, Bányiová, Bieliková and Piterková (2014) used DEA approach for corporate failure prediction of the agriculture sector. Their findings demonstrate aspects of application alternative DEA approach as a corporate prediction tool, and the ways of identification enterprises with high chance of potential bankruptcy.

Also Janová, Vavřina and Hampel (2012) also dealt with the issue of evaluating the financial health of companies through the DEA method, who discussed the possibility of application of recent results in DEA bankruptcy prediction models in a specific field of agribusiness. According to their results they concluded that the small test case study provides them with promising numbers. In their future research, further validation calculations using a larger dataset will be done.

In studies from Table 1, the authors have used different types of DEA models, and also have chosen different indicators as inputs or outputs to the model. For example, Paradi, Asmild and Simak (2004) used a different combination of total assets, working capital, earnings before income, tax, depreciation and amortization, retained earnings, shareholders equity, total current liabilities, interest expense, cash flow from operations, the stability of earnings and total liabilities as inputs and outputs. Feruś (2008) as inputs set daily return indicator and total liabilities indicator, and as outputs set net profit indicator, return on assets, return on equity and liquidity ratio. Mendelová and Bieliková (2017) chose the most frequently used indicators in the DEA model. As inputs were set liquidity 3, level of net working capital, liquidity from cash flow, self-financing coefficient, EBITDA revenue share, return on assets, operating return on sales and financial return on assets and as outputs set the maturity of short-term trade payables, total debt, long- term debt, cost ratio of economic activity

Name/s	Year	Country	Model	Aim of work	Results
Paradi, Asmild, Simak	2004	Kanada	BCC	To introduce the concept of worst practice DEA, which aims at identifying the worst performers by placing them on the frontier.	The results of the empirical application on credit risk evaluation validate the method.
Feruś	2008	Poland	basic DEA	A new procedure of forecasting credit risk to companies in the Polish economic environment.	The DEA method facilitates forecasting financial problems, including the bankruptcy of companies, in Polish economic conditions, and its efficiency is comparable to or even greater than that of the approaches implemented so far.
Sueyoshi, Goto	2009	USA	DEA-DA	Discuss methodological strengths and weaknesses of DEA and DEA–DA from the perspective of corporate failure.	DEA is a managerial tool for the initial assessment of corporate failure and DEA is useful for busy corporate leaders and financial managers; in contrast, DEA–DA is useful for researchers and individuals who are interested in the detailed assessment of bankruptcy and its failure process in a time horizon.
Mendelová, Bieliková	2017	Slovakia	ADD	To present a new proposal for diagnosing the corporate financial health by DEA, to predict financial distress of Slovak manufacturing companies using the proposed procedure, and to assess the potential of DEA as a tool for predicting financial distress of the company.	The application of the proposed procedure to Slovak manufacturing companies and its comparison with the logistic regression model and decision tree show relatively satisfactory results of the proposed methodology in terms of correct classification of non-bankrupt firms.
Xin, Hoe, Siew	2019	Malaysia	inverse-like DEA	Propose a method framework to estimate operational efficiencies and potential income gains considering the credit risk for banks.	The potential income gain can be estimated by the proposed inverse-like model credibly.
Horvathová, Mokrišová, Vrábliková	2019	Slovakia	CCR	To find out which financial indicators of the company are key performance indicators.	This research confirmed the possibility of integrating BSC and DEA.
Horvathová, Mokrišová, Vrábliková	2019	Slovakia	CCR	To find out which financial indicators of the company are key performance indicators.	This research confirmed the possibility of integrating BSC and DEA.

Source: prepared by authors

Table 1: Overview of the studies.

and interest cost ratio. As can be seen, all the mentioned authors used indicators that represent the main areas of company performance, namely, the area of liquidity, profitability, activity and indebtedness. In this paper, we will select indicators as inputs and outputs to the DEA model based on the procedure described in the following section.

Our contribution to the literature can be defined as follows. First, unlike most of the previous papers using one-year data, we focused on analysis ICT companies during the 2013-2017 period. Second, we do not apply the traditional models. However, we try to explore the possibilities of diagnosis corporate credit risk through Data Envelopment Analysis (DEA) method and design an appropriate model for the diagnosis of credit risk in the selected sample of companies.

Materials and methods

In this paper, we are going to use Data Envelopment Analysis method to evaluate the credit risk of selected companies. The selection of relative variables can be made based on a number of methods such as comparing box graphs, Mann-Whitney test, t-test, correlation analysis, or this selection can be made based on expert knowledge of the model creator or comparison of financial ratios between groups of companies in financial health and financial distress.

We have decided on the following procedure for determining inputs and outputs to the DEA model - inspired by Mendelová and Bieliková (2017):

1. multicollinearity elimination,
2. Mann-Whitney test,
3. find outliers,
4. distribution of indicators into DEA inputs and outputs.

According to Kassambara (2018), in multiple regression, two or more predictor variables might be correlated with each other. This situation is referred to as collinearity. Multicollinearity is an extreme situation, where collinearity exists between three or more variables, which means that there is redundancy between predictor variables. In the presence of multicollinearity, the result of the regression model is unstable. As mentioned by Kassambara (2018), multicollinearity for a given predictor can be assessed by computing a score by a variance inflation factor (VIF), which measures how much variance of the regression coefficient is inflated due to multicollinearity in the model.

Alin (2010) present the formula to calculate VIF for each variable:

$$VIF_i = \frac{1}{1-R_i^2} \quad \text{for i = 1, 2, …, k} \qquad (1)$$

R_i^2 is the coefficient of multiple determination of x_i on the remaining explanatory variables.

James et al. (2014) stated that the smallest possible value for VIF is 1, which indicates the complete absence of collinearity. Typically in practice, there is a small amount of collinearity among the predictors. As a rule of thumb, a VIF value that exceeds 5 or 10 indicates a problematic amount of collinearity.

The second step is the application of the Mann-Whitney test. The Mann-Whitney test, which is also known as the Wilcoxon rank-sum test, is a nonparametric test that allows two groups or conditions or treatments to be compared without making the assumption that values are normally distributed (Mann and Whitney, 1947; Wilcoxon, 1945).

Hart (2001) summarized the important points:

- The Mann-Whitney test is used as an alternative to a t-test when the data are not normally distributed.
- The test can detect differences in shape and spread as well as just differences in medians.
- Equally important differences in shape often accompany differences in population medians.
- Researches should describe the clinically important features of data and not just quote a P-value.

McKnight and Najab (2009) stated that the Mann-Whitney U is intended to determine if two groups (e.g. samples “a” and “b”) come from the same population (p), which is a null hypothesis significance test stipulating that both samples are subsets from the same population (i.e., H0: $(a,b) \subseteq p$). To test the null hypothesis, we first combine observations from two groups into a single group and rank the scores from 1 to N, where N is the total sample size ($n_a + n_b = N$). After ranking, the procedure divides the rank scores by group and computes a sum score for each group (T_a and T_b).

$$\begin{aligned} If\ n_a > n_b &: U = T_a - (n_a(n_a+1))/2 \\ If\ n_b > n_a &: U = T_b - (n_b(n_b+1))/2 \end{aligned} \qquad (2)$$

The U statistic has a discrete or uniform distribution that provides us with the ability to define a critical value, assign a probability to that value, and then test the null hypothesis. A critical value represents a probability level (typically .05). If the U statistic is greater than the critical value, then we reject the null hypothesis with the inference that both samples do not come from the same population.

Within the third step, we want to find out the outliers. According to Bogetoft and Otto (2011), outliers are firms that differ to a large extent from the rest of firms and therefore may end up being badly captured by the model or having too large an impact on the model. Outliers are often thought to be particularly troublesome to DEA because an outlier helps to span the frontier and may have a significant impact on the evaluation of several other firms.

The fourth step includes the distribution of indicators into DEA inputs and outputs. Financial indicators whose low value of the average indicates the company's financial problems in the future will be included in the group of inputs and financial indicators, whose high value of the average indicates the company's financial problems in the future, will be included in the group of outputs. In other words, the financial indicators that represent the financial strength and solvency of the company are selected for the inputs, and the financial indicators that represent the financial instability and insolvency of the company are chosen for the outputs.

After the mentioned steps, we apply the Data Envelopment Analysis to analyse the financial health of ICT companies. According to Mendelová and Bieliková (2017) construction, the production possibility curve consists of two steps:

- The production possibility curve is constructed on the basis of a whole sample of n companies. In this way, we identify businesses that create the production possibility curve. These companies will be considered to have a relatively high probability of their future financial problems.
- From the dataset, those businesses that formed the production possibility curve in the first step are omitted, and another production possibility curve is constructed again on such a reduced dataset. In this way, the negative impact of potential outliers is partially eliminated, and those businesses that create the production possibility curve at this stage are considered to have some probability of their future financial problems. However, this probability is relatively lower than that of those companies, that created the production possibility curve in the first step.

The authors used Altman´s (1968) identification of the resulting three zones in this model:

- Financial Distress Zone: Contains companies that created the production possibility curve in step 1, i.e. companies with a relatively high probability of future financial problems.
- Grey Zone: Contains businesses that created the production possibility curve in step 2, i.e., businesses with a lower probability of future financial problems.
- Financial Health Zone: Contains companies that did not create a financial distress frontier in any of the previous two steps, i.e., financially healthy businesses.

We have decided that the ADD model for the VRS condition (Charnes et al., 1985) will be applied to quantify the distances of companies from the production possibility curve because it is not necessary to select the input or output orientation of the model. ADD model for the VRS condition for company o, $o \in \{1,\dots, n\}$ is:

$$\max_{s^-,s^+,\lambda} e's^- + e's^+ \text{ under conditions:}$$

$$\begin{aligned} s^- &= x_0 - X\lambda, \\ s^+ &= Y\lambda - y_0, \\ \lambda &\geq 0, s^- \geq 0, s^+ \geq 0, \\ e'\lambda &= 1, \end{aligned} \tag{3}$$

where

n is the number of rated companies,

$X m\times n$ is a matrix of m inputs of n companies,

$Y s\times n$ is a matrix of s outputs of n companies,

e' is a row unit vector with all components equal to 1,

x_0 is column vector m inputs of the company o,

y_0 is column vector s outputs of the company o,

s^- is vector m slips of inputs of the company o,

s^+ is vector s slips of outputs of the company o,

$\lambda \in R^n$ is a vector of weights, which connect inputs and outputs.

Let (s^{-*}, s^{-*}, λ^*) be the optimal solution to the (3). Then the company o creates a frontier of financial distress in case the values of slips of inputs and outputs are zero, which means that $s^{-*} = 0$ and $s^{+*} = 0$.

To assess the competence of this model, we classified the total number of n rated enterprises into six groups.

Group A	companies in financial distress included in the financial distress zone
Group B	companies in financial distress included in the grey zone
Group C	companies in financial distress included in the financial health zone
Group D	companies in financial health included in the financial distress zone
Group E	companies in financial health included in the grey zone
Group F	companies in financial health included in the financial health zone

The companies classified in groups A and F can then be considered as correctly classified. The companies classified in groups B and E can be considered as neutrally classified, and the companies classified in groups C and D can be considered as misclassified. Altman (1968) stated that group D is Type I Error, and group C is Type II Error.

Numbers of companies classified into group i we marked as n_i, $i = A, B, C, D, E, F$.

Index of correct classification $I_{CC} \in [0,1]$ is

$$I_{CC} = \frac{n_A + n_F}{n} \quad (4)$$

index of neutral classification $I_{NC} \in [0,1]$ is

$$I_{NC} = \frac{n_B + n_E}{n} \quad (5)$$

and index of incorrect classification $I_{IC} \in [0,1]$ is

$$I_{IC} = \frac{n_C + n_D}{n} \quad (6)$$

Whereas $\sum_{i=A}^{F} n_i = n$, it must be true that $I_{CC} + I_{NC} + I_{IC} = 1$. The best situation is when values of I_{CC} are high, and values of I_{IC} are low.

The content of the following part of the article is an illustration and evaluation of the application of the above proposed DEA procedure for the case of Slovak ICT companies in the period 2013-2017. As it was mentioned in the introduction part ICT sector is considered to be specific, as it is currently the driving force behind the development and growth of the economy, and at the same time represents a tool for transforming the traditional functioning of the society. That is why it is necessary to study the financial health of these companies and this way also the probability of bankruptcy of these companies. In order to take into account, the potential differences that may exist between different sectors in the economy, only one sector of the economy was selected for analysis, the sector of Information services (SK NACE 63000). Due to the comparability of financial statements and subsequent quantification of financial indicators, only those companies that belonged to small, medium and large entities according to Slovak legislation were included in the analysis (micro-enterprises were not included due to differences in the structure of financial statements). We have used the Finstat´s dataset of financial indicators to get the list of companies that we are going to explore, as well as to obtain the necessary data to calculate the mentioned model.

Out of the total number of 69 indicators calculated separately for each company, we have selected 27 of them. The indicators that were not mentioned by the majority of companies (less than 50%) were removed from the database.

We have filtered the sample feed to include companies:

- not cancelled,
- earning more than 0€,
- which have available data for the 2013-2017 period.

A company in financial distress was considered to be a company that in next year met the criterion of defining a company in extension according to the valid legislation of the Slovak republic, i.e., the value of its payables exceeded the value of its assets, respectively the company reported a negative value of equity.

Results and discussion

All calculations, including quantification of the ADD DEA model, were performed in program R (specifically RStudio). We have used the following packages:

- faraway,
- readxl,
- deaR.

First of all, we have had to eliminate multicollinearity between indicators in order to exclude those indicators that provide relatively the same type of information. For this purpose, those indicators were excluded from the analysis for which a relatively high degree of correlation with other indicators was recorded – VIF ≤ 10.

Then we have performed the Mann-Whitney test of two independent selections to select those indicators whose values for healthy companies differ significantly from those for companies in bankruptcy according to equity – p-value less than .05.

As mentioned in the previous section, the occurrence of outliers can cause significant problems in the construction of the financial distress line. For this reason, our next step was to remove them. We have removed outsiders separately for healthy companies and separately for companies in bankruptcy according to equity because if we removed outliers regardless of whether they belong to the group of healthy or bankrupt companies according to equity, the analysis would remove bankrupt companies because those would consider them as outliers.

At the end of the pre-preparation of data for the DEA and the determination of inputs and outputs, we calculated the averages for the indicators that we had left. The values of these averages, as well as the results of the Mann-Whitney test, are shown in Table 2.

The result of the whole previous process is three inputs and four outputs, which we will use in the DEA analysis.

Before that, we have randomly selected 100 financially health companies and ten companies in financial distress according to equity for each year. The reason is that in this article, we want to compare the success of the DEA for the evaluation of the financial health of companies for five years and compare it with each other in this period.

We apply Data Envelopment Analysis as a tool for evaluating the financial health of companies. The success of diagnosing the financial health of Slovak ICT companies using the proposed two-step procedure is demonstrated in Table 3.

The overall financial situation of the company was assessed on a scale of three zones (financial health zone, grey zone, financial distress zone) with different degrees of the threat of financial distress over a one-year time horizon.

We have applied step 1 of the proposed model of Mendelová and Bieliková (2017), creating a production possibility curve that represents the financial distress zone and was made up of companies with the highest probability of future financial problems. Step 2 identifies the companies in the grey zone, which means that there is less probability of future financial problems.

The production possibility curve which represents financial distress in 2013 (2014, 2015, 2016, 2017) was formed by 7.27% (4.55%, 6.36%, 9.09%, 8.18%) of companies from the entire subset of 110 Slovak IT companies, which was calculated as

$$\begin{aligned}&percentage\ of\ companies\ that\ create\ the\ production\\&possibility\ curve\ in\ first\ step\\&=\frac{number\ of\ companies\ in\ group\ A+group\ D}{total\ number\ of\ companies}*100\end{aligned} \tag{7}$$

Group D includes companies that are on the curve of financial distress but are, in fact, financially healthy. However, this fact does not represent a significant problem from the point of view of the application, as the result of the diagnostics caused by the increased interest in the financial situation of the company does not bring negative consequences for the interested parties of the company. The production possibility curve in the second step (grey zone) in 2013 (2014, 2015, 2016, 2017) was formed by 19.09% (10.91%, 9.09%, 6.36%, 9.09%) of companies, which was calculated as

$$\begin{aligned}&percentage\ of\ companies\ that\ create\ the\ production\\&possibility\ curve\ in\ second\ step\\&=\frac{number\ of\ companies\ in\ group\ B+group\ E}{total\ number\ of\ companies}*100\end{aligned} \tag{8}$$

Indicators	Mann-Whitney test (p-value)	Average of financially health companies	Average of companies in financial distress	Input / Output
Net debt	0.00000	-250723.11	-37258.57	output
Total insolvency	0.00002	0.92	1.29	output
Return on long-term capital (EBIT)	0.00000	0.40	0.13	input
Repayment period of liabilities	0.00679	152.05	173.49	output
Repayment period of liabilities in relation to sales	0.00137	73.19	81.62	output
Effective tax rate	0.00000	0.25	0.19	input
Coverage of personnel costs and depreciation	0.00000	1.29	1.04	input

Source: Own calculations

Table 2: Inputs and outputs.

This second curve represents a zone with a lower degree of risk and represents the possibility of financial problems in the following year. The model correctly included up to 73.64% (84.55%, 84.55%, 84.55%, 82.73%) of financially healthy companies in the financial health zone, which was calculated as

$$\begin{aligned}&\textit{percentage of companies that create the production}\\&\textit{possibility curve in second step}\\&= \frac{\textit{number of companies in group } C + \textit{group } F}{\textit{total number of companies}} * 100\end{aligned} \qquad (9)$$

According to this model, these companies are outside the threat of financial distress.

	2013	2014	2015	2016	2017
A	3	1	2	4	3
	2.73%	0.91%	1.82%	3.64%	2.73%
B	6	2	2	1	2
	5.45%	1.82%	1.82%	0.91%	1.82%
C	1	7	6	5	5
	0.91%	6.36%	5.54%	4.55%	4.55%
D	5	4	5	6	6
	4.55%	3.64%	4.55%	5.45%	5.45%
E	15	10	8	6	8
	13.64%	9.09%	7.27%	5.45%	7.27%
F	80	86	87	88	86
	72.73%	78.18%	79.09%	80.00%	78.18%
∑	110	110	110	110	110
I_{CC}	75.45%	79.09%	80.09%	83.64%	80.91%
I_{NC}	19.09%	10.91%	9.09%	6.36%	9.09%
I_{IC}	5.45%	10.00%	10.00%	10.00%	10.00%

Source: Own calculations

Table 3: Results of diagnosing the financial health of Slovak ICT companies.

To assess the overall classification ability of the model can be evaluated based on the mentioned indices, the index of correct classification, the index of neutral classification and the index of incorrect classification.

In 2013 (2014, 2015, 2016, 2017) the model correctly classified 83 (87, 89, 92, 89) companies from the total set of 110 companies, which in a percentage expression represents 75.45% (79.09%, 80.09%, 83.64%, 80.91 %). The index of correct classification achieves relatively satisfactory results in all years. Using the parameters shown in Table 2 to the model, DEA showed a relatively high percentage. Therefore, we propose the use of these indicators to analyse the financial health of IT companies in the long term using the DEA method.

The index of neutral classification indicates the ratio of companies that were included in the grey zone. Companies in the grey zone are characterized by a certain degree of risk of over-indebtedness, but lower compared to the financial distress zone. The value of the index of neutral classification ranges from 6.36% to 19.09% in the observed years. It can be stated that in the grey zone, financially healthy companies significantly prevail over companies in financial distress for all years.

The incorrect classification index reached 5.45% in 2013 and in the other years (2014, 2015, 2016, 2017) it reached the level of 10%, which means that the model in 2013 (2014, 2015, 2016, 2017) incorrectly classified six companies in 2013 and 11 companies in other years.

Delina and Packová (2013) in their study dealt with the validation of selected models (Altman's Z-score, Beerman's discriminant function - Bonity Index and IN05 Index) on real data of companies established in Slovakia in the period 1993-2007, while they developed new modified model while using regression analysis to get higher predictive performance on analysed sample than chosen models. Compared to the traditional models used in this study, the error rate of the DEA model we used is not so high (The Altmans Z-score was 87.62%, the Bonity Index was 78.02% and the IN05 was 85.41%), and therefore we can consider the DEA model to be suitable for application in the conditions of ICT companies in Slovakia.

At first sight, the results of this model are satisfactory, but if we look at the number of companies included in group A (enterprises in financial distress classified as financial distress) so we see that out of the total number of 10 companies classified in the financial distress zone according to equity, in 2013 (2014, 2015, 2016, 2017) only 3 (1, 2, 4, 3) companies were included in this category, which in neither of these years nor 50% of the total number of companies in financial distress. Such a relatively high error rate of the model is considered a high risk in terms of possible consequences and potential costs of neglecting financial distress signals.

As mentioned in the literature review chapter, it is crucial for a company which wants to apply for a loan to assess its financial situation and be financially healthy (low credit risk). If it is located in the grey zone (greater credit risk), it should take steps not to be at risk of failure and to be included in the financial health zone in the next period.

The worst option is if the company is included in the financial default zone because at that time it poses a significant credit risk for the bank, which is why the bank does not approve the loan.

Conclusion

Data Envelopment Analysis is increasingly being used in various areas of finance, health sector, agricultural sector and others. This article dealt with its use in the evaluation of the financial health of a selected sample of companies in the period 2013 to 2017, and consequently, the credit risk they may pose to financial entities. Mainly, we have applied the ADD DEA model and evaluated the classification indexes of this model. As the main contribution of the work, we consider the proposal of specific financial indicators for the analysis of the financial health of companies in the ICT sector in a multi-year period using the DEA model. We justify the choice of these indicators by the relatively high success of the DEA model, in which these indicators were used as inputs and outputs. The paper can also serve as a methodological approach to evaluating the financial health of companies through the DEA model, regardless of the industry. The paper can also serve as additional teaching material on the issue of using DEA to evaluate the financial health of the company.

The results show that DEA achieves a satisfactory value of a correct classification into the relevant zone (financial health zone, grey zone, financial distress zone). On the other hand, the relatively high error rate of the DEA model in the identification of companies in financial distress encourages the discussion of the adequacy and appropriateness of using DEA in diagnosing the financial health of companies in real conditions. We believe that one of the possible factors that could negatively affect the classification accuracy of the used DEA model is the structure of the data set. In the years, which we have analysed, the ratio between financially healthy companies and companies in financial distress was 10: 1. This low 9% ratio of financially distressed companies in the dataset may have led the DEA to identify with its error rate a small percentage of the total number of financially distressed companies included in the analysis. In our opinion, this was the fact that greatly influenced the results and seemingly degraded the ability of the DEA model used. We assume that a higher ratio of companies in financial distress in the data set will, on the contrary, favour the classification capacity of the DEA, so that the DEA will be able to correctly classify a higher percentage of companies in the financial distress zone.

The method applied by us is suitable for application in other sectors of the national economy. An example is the agro-sector, where this method can also be used by agro-policy makers in evaluating the financial health of companies, which is part of the decision to provide funds. It is on the basis of the use of this method that it is possible to achieve that the provided state or EU funds are spent efficiently.

It mentioned perspective regarding the increase of the classification ability of the proposed DEA model as well as the comparison of this alternative method with the traditional models used to evaluate the financial health of the company, such as Altman's Z-score, Zmijewski's X-score, IN 05 or Quick test for different industries, are the basis for our further research in the future. We also want to research whether there is a link between the performance at the operational (cost-oriented) and financial (profit-oriented) spaces of the Slovak company as described by Tsolas (2015) on Greek companies. The aim will be to determine whether the company's profit is the driving force of its success in the industry and to analyse the operating performance efficiency of Slovak companies in the sample but also to assess their performance in earnings and cash flow generation.

Acknowledgements

This paper was supported by VEGA project No. 1/0794/18.

Corresponding authors
Assoc. prof. Kristína Kočišová, PhD.
Department of Banking and Investments, Faculty of Economics, Technical University of Košice
Němcovej 32, 040 01 Košice, Slovak Republic
E-mail: kristina.kocisova@tuke.sk
ORCID: 0000-0003-0784-441X

References

[1] AlAli, M., AlShamali, M., AlAwadhi, K. and AlSabah, A. (2018) "The use of Zmijewski model in examining the financial soundness of oil and gas companies listed at Kuwait Stock Exchange“, *International Journal of Economics*, Commerce and Management Research Studies, Vol. 1, No. 2, pp. 15-21. ISSN 2456-2165.

[2] Alin, A. (2010) "Multicollinearity“, *Wiley Interdisciplinary Reviews: Computational Statistics,* Vol. 2010, No. 2-3., pp. 370-374. ISSN 1939-5108. DOI 10.1002/wics.84.

[3] Altman, E. I. (1968) "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy“, *Journal of Finance*, Vol. 23, No. 4, pp. 589-609. ISSN 00221082. DOI 10.2307/2978933.

[4] Aminian, A., Mousazade, H. and Khoshkho, O. I. (2016) "Investigate the ability of bankruptcy prediction models of Altman and Springate and Zmijewski and Grover in Tehran Stock Exchange“, *Mediterranean Journal of Social Sciences*, Vol. 7, No. 4 S1, pp. 208. ISSN 2039-2117. DOI 10.5901/mjss.2016.v7n4S1p208.

[5] Bandyopadhyay, A. (2007) "Credit risk models for managing bank’s agricultural loan portfolio“, Munich Personal RePEc Archive.

[6] Banker, R. D., Charnes, A. and Cooper, W.W. (1984) "Some Models for Estimating Technical Scale Inefficiencies in Data Envelopment Analysis“, *Management Scienc*e, Vol. 30, No. 9, pp. 1078-1092. ISSN 1526-5501. DOI 10.1287/mnsc.30.9.1078.

[7] Bányiová, T., Bieliková, T. and Piterková, A. (2014) "Prediction of agricultural enterprises distress using data envelopment analysis“, *Proceedings of the 11th International Scientific Conference European Financial Systems*, Masaryk University, Brno, pp. 18-25. ISBN: 978-80-210-7153-7.

[8] Beaver, W. H. (1966) "Financial ratios as predictors of failure“, *Journal of Accounting Research*, Vol. 30, No. 9., pp. 71-111. E-ISSN 1526-5501, ISSN 0025-1909. DOI 10.2307/2490171.

[9] Boďa, M. and Úradníček, V. (2019) "Predicting Financial Distress of Slovak Agricultural Enterprises“, *Ekonomický časopis*, Vol. 67, No. 4, pp. 426-452. ISSN 0013-3035.

[10] Bogetoft, P. and Otto, L. (2011) "Additional Topics in DEA", In: *Benchmarking with DEA, SFA, and R. International Series in Operations Research & Management Science*, Vol. 157, Springer, New York, NY. ISBN 978-1-4419-7960-5, E-ISBN 978-1-4419-7961-2. DOI 10.1007/978-1-4419-7961-2_5.

[11] Brewer, B. E., Wilson, C. A., Featherstone, A. M., Harris, J. M., Erickson, K. and Hallahan, C. (2012) "Measuring the financial health of US production agriculture“, *Journal of ASFMRA*, pp. 178-193. ISSN 0003116X.

[12] Delina, R. and Packová, M. (2013) "Validácia predikčnýchch bankrotových modelov v podmienkach SR“, *E+ M Ekonomie a management,* Vol. 2013, No. 3, p. 101-112. ISSN 2336-5604. (in Slovak).

[13] Ékes, K. S. and Koloszár, L. (2014) "The efficiency of bankruptcy forecast models in the Hungarian SME sector“, *Journal of Competitiveness*, Vol. 6, No. 2, pp. 56-73. ISSN 1804-1728. DOI 10.7441/joc.2014.02.05.

[14] Fakhri Husein, M. and Tri Pambekti, G. (2014) "Precision of the models of Altman, Springate, Zmijewski, and Grover for predicting the financial distress“, *Journal of Economics, Business, and Accountancy Ventura*, Vol. 17, No. 3, pp. 405-416. ISSN 2087-3735. DOI 10.14414/jebav.14.1703010.

[15] Feruś, A. (2008) "The DEA method in managing the credit risk of companies“, *Ekonomika*,Vilnjus University Press, E-ISSN 1392-1258, ISSN 2424-6166 Vol. 2008, No. 84. [Online]. Available: http://www.journals.vu.lt/ekonomika/article/view/5126/3383. ISSN 1392-1258. [Accessed: 4 Jan. 2020].

[16] Grover, J. and Lavin, A. (2001) "Financial Ratios, Discriminant Analysis and The Prediction of Corporate Bankruptcy: a Service Industry Extension of Altman’s Z-Score Model of Bankruptcy Prediction“, Working Paper. Southern Finance Assosiation Annual Meeting.

[17] Hart, A. (2001) "Mann-Whitney test is not just a test of medians: differences in spread can be important", *BMJ*, Vol. 323, No. 7309, pp. 391-393. ISSN 0959-8138, 1756-1833, 0959-8138, 1756-1833, 0007-1447, 0267-0623. DOI 10.1136/bmj.323.7309.391.

[18] Horváthová, J., Mokrišová, M. and Vrábliková, M. (2019) "Integration of balanced scorecard and data envelopment analysis to measure and improve business performance", *Management Science Letters*, Vol. 9, No. 9. ISSN 1321-1340. DOI 10.5267/j.msl.2019.5.017.

[19] Charnes, A., Cooper, W. W. and Rhodes, E. (1978) "Measuring the Efficiency of Decision Making Units", *European Journal of Operational Research*, Vol. 2, No. 6, pp. 429-444. ISSN 0377-2217. DOI 10.1016/0377-2217(78)90138-8.

[20] Charnes, A., Cooper, W. W., Golany, B., Seiford, L. M. and Stutz, J. (1985) "Foundations of Data Envelopment Analysis for Pareto-Koopmans Efficient Empirical Production Functions", *Journal of Econometrics*, Vol. 30, No. 12, pp. 91-127. ISSN 0304-4076. DOI 10.1016/0304-4076(85)90133-2.

[21] Chaudhuri, A. and Ghosh, S. K. (2017) "*Bankruptcy prediction through soft computing based deep learning technique*", Springer. ISBN 978-981-10-6683-2.

[22] James, G., Witten, D., Hastie, T. and Tibshirani, R. (2013) "*An introduction to statistical learning*", New York: Springer, Vol. 112, pp. 3-7. ISBN 978-1-4614-7137-0.

[23] Janová, J., Vavřina, J. and Hampel, D. (2012) "DEA as a tool for bankruptcy assessment: the agribusiness case study", *Proceedings of the 30th International Conference Mathematical Methods in Economics*, Vol. 2012, pp. 379-383. ISBN 978-80-7248-779-0.

[24] Jedik, A. and Stalgienė, A. (2018) "The likelihood of farms bankrupcy: The case of Lithuanian family farms", *Management Theory and Studies for Rural Business and Infrastructure Development*, Vol. 40, No. 2, pp. 198-205. E-ISSN 2345-0355. DOI 10.15544/mts.2018.19.

[25] Káčer, M., Ochotnický, P. and Alexy, M. (2019) "The Altman's Revised Z'-Score Model, Non-financial Information and Macroeconomic Variables: Case of Slovak SMEs", *Ekonomický časopis*, Vol. 67, No. 4, pp. 335-366. ISSN 0013-3035.

[26] Karas, M., Reznakova, M. and Pokorny, P. (2017) "Predicting bankruptcy of agriculture companies: Validating selected models", *Polish Journal of Management Studies*, Vol. 15, No. 1. ISSN 2081-7452. DOI 10.17512/pjms.2017.15.1.11.

[27] Kassambara, A. (2018) "*Machine Learning Essentials: Practical Guide in R*", STHDA, p. 209. ISBN 9781986406857.

[28] Kiaupaite-Grushniene, V. (2016, December) "Altman Z-score model for bankruptcy forecasting of the listed Lithuanian agricultural companies", *5th International Conference on Accounting, Auditing, and Taxation (ICAAT 2016)*, Atlantis Press. ISBN 978-94-6252-261-9. ISSN 2352-5428. DOI 10.2991/icaat-16.2016.23.

[29] Kotulič, R., Király, P. and Rajčániová, M. (2007) "*Finančná analýza podniku*", Bratislava, Iura Edition, p. 206. ISBN 978-80-8078-117-0. (in Slovak).

[30] Kralicek, P. (1993) "Základy finančního hospodaření", Prague, Linde, p. 110, ISBN 80-85647. (in Czech).

[31] Krausová, A. (2018) "Abuse of market power in ICT sector", *The Lawyer Quarterly*, Vol. 8, No. 1, pp. 78-81. ISSN 1805-840X.

[32] Mann, H. B. and Whitney, D. R. (1947) "On a test of whether one of two random variables is stochastically larger than the other", T*he Annals of Mathematical Statistics,* Vol. 18, No. 1, pp. 50–60. ISSN 00034851.

[33] McKnight, P. E. and Najab, J. (2010) "Mann-Whitney U Test", *In: The Corsini Encyclopedia of Psychology*. E-ISBN 9780470479216, ISBN 9780470170243. DOI 10.1002/9780470479216.

[34] Mendelová, V. and Bieliková, T. (2017) "Diagnostikovanie finančného zdravia podnikov pomocou metódy DEA: Aplikácia na podniky v Slovenskej republike", *Politická ekonomie*, Vol. 2017, No. 1, pp. 26-44. ISSN 23368225. DOI 10.18267/j.polek.1125. (in Slovak)

[35] Paradi, J. C., Asmild, M. and Simak, P. C. (2004) "Using DEA and Worst Practice DEA in Credit Risk Evaluation", *Journal of Productivity Analysis*, Vol. 21, No. 2, pp. 153-165. ISSN 0895562X. DOI 10.1023/B:PROD.0000016870.47060.0b.

[36] Sueyoshi, T. and Goto, M. (2009) "Methodological Comparison Between DEA (Data Envelopment Analysis) and DEA-DA (Discriminant Analysis) from the Perspective of Bankruptcy Assessment", *European Journal of Operational Research*, Vol. 199, No. 2, pp. 561-575. ISSN 03772217. DOI 10.1016/j.ejor.2008.11.030.

[37] Tone, K. (2001) "A Slacks-Based Measure of Efficiency in Data Envelopment Analysis", *European Journal of Operational Research*, Vol. 130, No. 3, pp. 498-509. ISSN 03772217. DOI 10.1016/s0377-2217(99)00407-5.

[38] Vaněk, J., Jarolímek, J. and Vogeltanzová, T. (2011) "Information and Communication Technologies for Regional Development in the Czech Republic – Broadband Connectivity in Rural Areas", *Agris on-line Papers in Economics and Informatics*, Vol. 3, No. 3, pp. 67-76. ISSN 1804-1930.

[39] Vavřina, J., Hampel, D., and Janová, J. (2013) "New approaches for the financial distress classification in agribusiness", *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, Vol. 61, No. 4, pp. 1177-1182. E-ISSN 2464-8310, ISSN 1211-8516. DOI 10.11118/actaun201361041177.

[40] Wilcoxon, F. (1945) "Individual comparisons by ranking methods", In: *Breakthroughs in Statistics*, pp. 196-202. ISBN 978-1-4612-4380-9. DOI 10.1007/978-1-4612-4380-9_16.

[41] Xin, A. L. J., Hoe, L. W. and Siew, L. W. (2019) "An empirical evaluation on the credit risk efficiency of financial institutions in Malaysia with data envelopment analysis model", *AIP Conference Proceedings*, Vol. 2138, No. 1. DOI 10.1063/1.5121102.

[42] Zhao, J., Barry, P. J. and Katchova, A. L. (2008) "Signaling credit risk in agriculture: implications for capital structure analysis", *Journal of Agricultural and Applied Economics*, Vol. 40, No. 3. pp. 805-820. ISSN 1074-0708. DOI 10.1017/S1074070800002340.

Volume XII Number 4, 2020

Open Source Framework for Enabling HPC and Cloud Geoprocessing Services

José Miguel Montañana[1], Paolo Marangio[2], Antonio Hervás[3]

[1] High Performance Computing Center Stuttgart (HLRS), University of Stuttgart, Germany

[2] ATOS Research and Innovation (ARI), AI, Data & Robotics Unit, Madrid, Spain

[3] Inst. Matemática Multidisciplinar (IMM), Universitat Politècnica de València, Spain

Abstract

Geoprocessing is a set of tools that can be used to efficiently address several pressing chal-lenges for the global economy ranging from agricultural productivity, the design of transport networks, to the prediction of climate change and natural disasters. This paper describes an Open Source Framework developed, within three European projects, for Ena-bling High-Performance Computing (HPC) and Cloud geoprocessing services applied to agricultural challenges. The main goals of the European Union projects EUXDAT (EUro-pean e-infrastructure for eXtreme Data Analytics in sustainable developmenT), CYBELE (fostering precision agriculture and livestock farming through secure access to large-scale HPC-enabled virtual industrial experimentation environment empowering scalable big data analytics), and EOPEN (opEn interOperable Platform for unified access and analysis of Earth observatioN data) are to enable the use of large HPC systems, as well as big data management, user-friendly access and visualization of results. In addition, these projects focus on the development of software frameworks, and fuse Earth-observation data, such as Copernicus data, with non-Earth-observation data, such as weather, environmental and social media information. In this paper, we describe the *agroclimatic-zones* pilot used to validate the framework. Finally, performance metrics collected during the execution (up to 182 times speedup with 256 MPI processes) of the pilot are presented.

Keywords

High performance computing, cloud computing, big data; agriculture, land monitoring, geoprocessing.

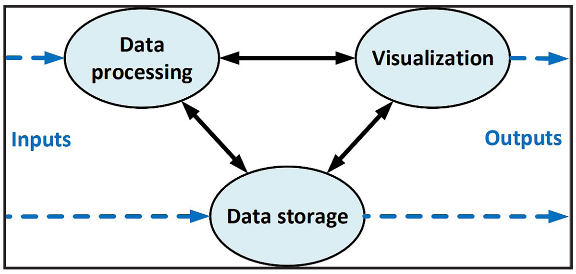
Montañana, J. M., Marangio, P. and Hervás, A. (2020) "Open Source Framework for Enabling HPC and Cloud Geoprocessing Services", *AGRIS on-line Papers in Economics and Informatics*, Vol. 12, No. 4, pp. 61-76. ISSN 1804-1930. DOI 10.7160/aol.2020.120405.

Introduction

Geoprocessing is a set of tools, generally intended for the mathematical processing carried out by a Geographic Information System (GIS). These tools consist of es-sentially three parts, as shown in Figure 1, namely data storage, computational pro-cessing, and visualization or access to re-sults.

During the last decades, the results of geo-processing have greatly improved thanks to the exponential technological progress in computational power. However, improving the efficiency of agricultural productivity requires solving the techno-logical challenge of increasing both the amount of data to be stored and the com-putational load by several orders of magnitude.

In order to tackle these challenges, during the last decades, researchers and professionals in the area have worked towards im-proving overall code performance in several ways, including parallelization of code libraries, structuring of the data, as well as balancing the computational load in clusters of computers (Figure 1).



Source: own research and processing

Figure 1: Fundamental components of geoprocessing systems and their interrelationship.

As a result, MPI and OpenMP now represent some of the most popular tools for code paralleliza-tion. And more recently, cloud computing and High-Performance Computing (HPC) have become the standard for Big Data processing. In particular, HPC systems are currently able to provide the best computing performance as well as enhanced data sharing between computing nodes (Mi-neter et al., 2000; Zhang, 2010; Li, 2020).

In the next sections, we define some theoretical and practical concepts that need to be considered for an efficient use of HPC systems.

Hardware for HPC

One important difference between HPC and Cloud systems is their interconnection network. Most modern HPC systems are clusters of Symmetric Multi-Processing (SMP) nodes with high-speed interconnection network, which eases the collaborative computation between nodes as well as the sharing of data between them. On the other hand, Cloud computing nodes have lower performance interconnection networks than HPC. Therefore, the parallelized applications running in Cloud should have less communication between nodes in order to not lose performance.

A SMP node consists of multiple identical processing elements, with identical memory access. The memory inside of the nodes allows strongly coupled processing and communication. The computation carried out among multiple nodes will have higher communication latency between cores when they are on different nodes and higher memory access latency when the data required by one node is stored in the memory of another node.

Parallelization Strategies for HPC

There are two main resources that can be distributed: the processing elements and the memory. Considering this, the parallelization of an hypothetical application will consists in deciding how to distribute the computational load among the processing elements, and how to distribute the data when using more than one node (case of distributed memory).

The distribution of the computation load and data requires defining how the internode communication will be performed, which is a very important aspect to take into account. Inefficient communication can make the processing elements stay idle while waiting for data from other processing elements or memory, which will produce an inefficient use of processing elements and therefore extend the execution time. Montañana (2010) provides more details on HPC architecture and interconnection networks.

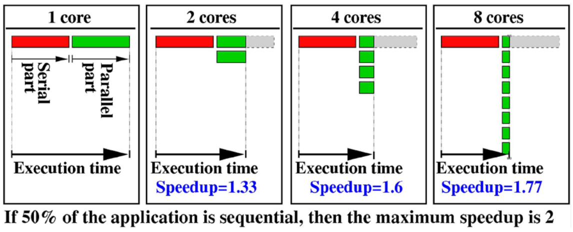
MPI and OpenMP are some of the most popular libraries and tools used for HPC systems. MPI is typically used for distributed memory systems.

Parallelization performance

A hypothetical application can be divided into two parts, namely the part that needs to be executed sequentially (s), and the part that can run in parallel (p). Equation 1 represents the decimal fraction of that distribution, where 1 stands for the total, and s and p values are between 1 and 0:

$$1 = s + p \tag{1}$$

Figure 2 describes the above concept pictorially, by showing an application that exhibits the potential of running 50% of its code in parallel (green bar in the figure). The figure shows that the total execution time becomes smaller as the number of used cores is increased. This is because the part of the code that runs in parallel is split evenly among the processors that execute their respective portion of the code concurrently.



Source: own representation of the Amdahl's law (Amdahl, 1967)

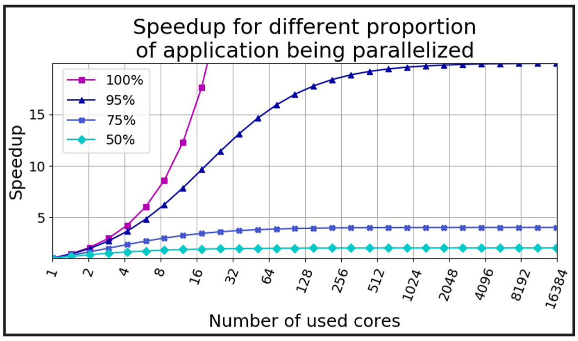
Figure 2: Reduction of execution time and increase of Speedup with the increase of the number of cores used.

The figure also illustrates the concept of *Speedup* (in latency), which corresponds to the ratio between the execution time sampled when the application is executed using only one core, and when using an increased number of them. Calculating the Speedup is useful for determining the im-pact of specific code changes on perfor-mance. Equation 2 shows how the Speedup is calculated as the division of Equation 1 for one and n processes, where n is the number of cores used, or the number of parallel executions of the part of the appli-cation that can be parallelized. This equa-tion is a formulation of Amdahl's law (Amdahl, 1967).

$$Speedup = \frac{s + \frac{p}{1}}{s + \frac{p}{n}} = \frac{1}{(1-p) + \frac{p}{n}} \tag{2}$$

From this, it follows that the execution time of a parallelized application is given by the execution time required to run the sequential part of the application plus the execution time of the parallelized part. This also means that the proportion of the application that can be executed in parallel dictates the theoretical maximum *Speedup* that can be achieved by a given parallel application.

Figure 3 shows the *Speedup* for different proportions of an application that can be parallelized. The ideal case where 100% of an application is parallelized is virtually impossible to attain due to software and hardware limitations, such as communication latencies and parallelization overheads.



Source: own processing of the Amdahl's law (Amdahl, 1967)

Figure 3: Speedup for different proportion of application being parallelized. The slopes of the curves are reduced from the point where the improvements no longer have a significant impact on the total run time. The curves saturate at the level where practically all the execution time is due to the non-parallelized part of the code.

Our experience has taught us that achieving performance improvement is relatively simple when the proportion of code being parallelized is small and the parallel code is executed on a small number of cores. However, the improvement becomes more difficult to achieve as the number of cores required for a given parallelization strategy increases. For this reason, there is a trade-off between the effort of parallelization and the performance improvements that can be obtained through parallelization. The most frequently used applications should also be the ones that receive the most parallelization efforts.

Contributions of the paper and overview of EU projects

In this paper, we significantly extend on our previous description of the Open-Source framework for enabling HPC and Cloud geoprocessing services (Montañana et al, 2020b). We will also show some preliminary results obtained by running the *agroclimatic-zones* pilot within the framework in the supercomputer Hawk at High Performance Computing Center Stuttgart (HLRS).

The main contribution of this paper is to showcase the latest developments in the creation of innovative platforms that solve several technological challenges that are relevant to geoprocessing, such as the integration of data from different origins and formats, the definition of interfaces for geoprocessing applications, and the capability of executing such applications on modern computing solutions like HPC and in the Cloud. Addressing these challenges requires a consideration of additional aspects such as allowing larger data transfer, and the enforcement of secure access and control of the data and the computational results.

All these challenges are of definite relevance to the EU projects EUXDAT (EUXDAT, 2020), EOPEN (VEOPEN, 2020), and CYBELE (Davy et al., 2020). In fact, these projects focus on developing solutions for the collection of big data from different sources, data transfer to large-scale HPC and Cloud Computing infrastructures for processing, the development of visualization tools as well as secure access to computational results.

In the next subsections, we provide a summary of the goals of the three projects:

EUXDAT

EXUDAT proposes an e-Infrastructure for enabling Large Data Analytics-as-a-Service, which addresses the problems related to the current and future huge amount of heterogeneous data to be managed and processed within the agricultural domain. EUXDAT builds on existing mature components by providing an advanced frontend, where users develop applications on top of an infrastructure based on HPC and Cloud. The frontend provides monitoring information, visualization, different distributed data analytic tools, enhanced data and processes catalogs. EUXDAT includes a large set of data connectors such as Unmanned Aerial Vehicles (UAVs), Copernicus, and field sensors for scalable analytics. Figure 4 shows the type of field sensors deployed for the EUXDAT project in farming areas. These instruments (Pessl, 2020) allow collection of a wide range of data about remote are-as, such as depth of precipitation, air tem-perature, air humidity, global radiation, wind speed, soil temperature, and leaf wetness.

Source: own processing, pictures took by authors

Figure 4: Example of a typical field sensor deployed in farming areas proposed within EUXDAT project. The sensor shown is an iMETOS 3.3 data logger developed by Pessl Instruments GmbH.

As for the brokering infrastructure, EUXDAT aims at optimizing data and resource usage. In addition to a mecha-nism for supporting data management linked to data quality evaluation, EUXDAT proposes a method to orches-trate the execution of tasks that is able to identify whether the best target for execut-ing a given application is HPC or Cloud. It uses monitoring and profiling information for making decisions based on trade-offs related to cost, data constraints, efficiency, and resource availability. During the pro-ject, EUXDAT is in contact with scientific communities, in order to identify new trends and datasets, for guiding the evolu-tion of the e-Infrastructure. The result of the project will be an integrated e-Infrastructure that encourages end-users to create new applications for sustainable development.

EUXDAT demonstrates real agriculture scenarios, land monitoring, and energy efficiency for sustainable development, as a way to support planning policies.

CYBELE

CYBELE is a European research project combining Agriculture, HPC, and Big Da-ta. It involves 31 research institutes and enterprises across EU countries. It stands for "*Fostering Precision Agriculture and Livestock Farming through Secure Access to Large-Scale HPC-Enabled Virtual Industrial Experimentation Environment Empowering Scalable Big Data Analytics*" (Perakis, 2020).

CYBELE generates innovation and creates value in the domain of agri-food, and its verticals in the sub-domains of Precision Agriculture (PA) and Precision Livestock Farming (PLF) specifically, as demonstrated by the real-life industrial cases to be supported, empower capacity building within the industrial and research community. The project aspires at demonstrating how the convergence of HPC, Big Data, Cloud Computing and the Internet of Things (IoT) can revolutionize farming, reduce food scarcity, increase food supply, bringing social, economic and environ-mental benefits. It develops large scale HPC-enabled testbeds and delivers a distributed big data management architecture and a data management strategy.

EOPEN

The objective of EOPEN is to fuse Earth Observation (EO) data with multiple, heterogeneous, and big data sources, in order to improve the monitoring capabilities of the future EO downstream sector. EO data consists of the Copernicus and Sentinel data, while the non-EO data is weather, environmental, and social media infor-mation.

The fusion between these diverse types of data is carried out at the semantic level, to provide reasoning mechanisms and in-teroperable solutions, through the seman-tic linking of information. The processing of large streams of data is based on open-source and scalable algorithms in change detection, event detection, data clustering, which are built on HPC infrastructures.

Alongside this enhanced data fusion ap-proach, an innovative architecture over-arching Joint Decision & Information Governance is combined with the tech-nical solution to assist with decision making and visual analytics. EOPEN is demonstrated through real use case scenarios in flood risk monitoring, food security, and climate change monitoring.

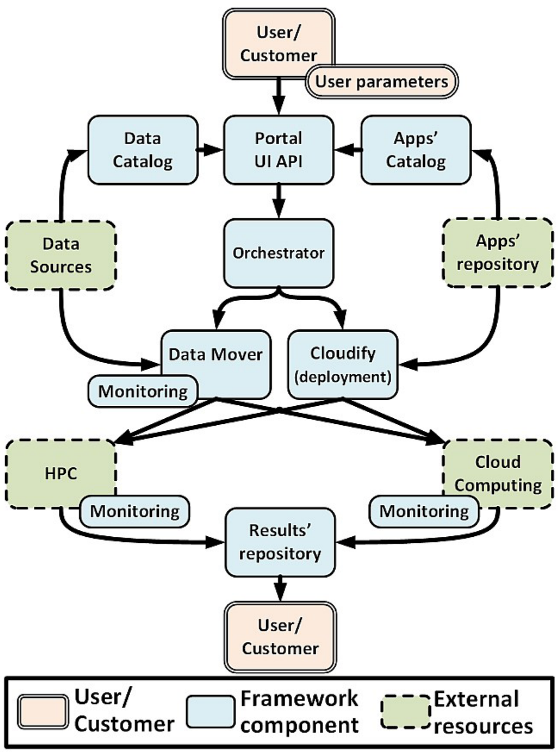
Materials and methods

Platform implementation

The common goal shared across these projects is to develop a sustainable approach that facilitates access to data, geoprocessing applications, state-of-the-art solutions for big-data management, as well as computational resources provided by Cloud platforms and HPC centers.

Therefore, the target of the implementation of the platform is to provide an open source system that can be used on HPC and Cloud computing systems for hosting commercial applications or products, as well as to ensure continuity in the use of a given platform after the respective EU project has been finalized. For example, although it may not be apparent, the reason for including support for accounting and billing in such platforms is to facilitate the future reuse of code. In fact, in order to ensure that a given piece of code can be successfully reused later in time, the costs of using large computer systems, as well as the cost of data acquisition from proprietary sources must be already considered during the development stage.

Figure 5 illustrates the main components of a typical infrastructure platform as well as their interrelationship. Each of the components in a given infrastructure platform has a clearly defined User Interface (UI).



Source: own research and processing

Figure 5: Representative infrastructure platform for the three EU projects described in the paper.

Portal UI API

The first component encountered by the user is the Portal User Interface (UI) Application Programming Interface (API). The portal UI API provides users with a list of available applications and the data catalog available for them. This component supports the development of applications such as mobile devices or web interfaces while abstracting away the complexity of the other components. For example, the user does not need to consider neither the complexity or format of the data, nor the different data sources, because it is encapsulated by the platform internally. Thus, once the user selects the task to perform, such as the prediction of temperature for a particular land area on a particu-lar date, the user just waits for the result. Needless to say, the time needed until a response is received is reduced by multiple orders of magnitude when using a large-scale HPC system.

Data Catalog

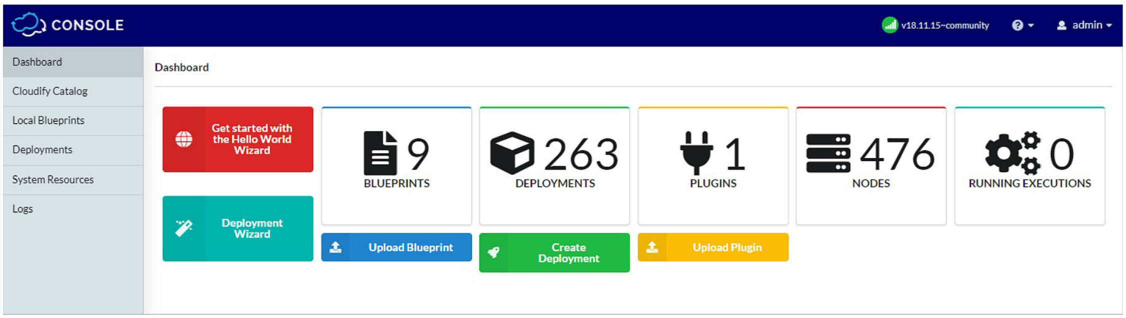
The data catalog collects data from different data sources, which may or not be free of charge. Similarly, the catalog of applications may include free applications such as those developed and hosted within the platform of a given project, or commercial applications. This is done in order to ad-vertise the platform to third parties that may wish to use its services in order to commercialize applications or data.

Orchestrator

The user request is submitted to the orchestrator, which is responsible for the transfer and execution of the applications on the computing resources. Based on the user request, the orchestrator selects the appropriate computational resource (i.e. HPC or Cloud) onto which to execute the task. Orchestrators such as Cloudify typically use an application model Domain Specific Language (DSL) based on Topology and Orchestration Specification for Cloud Applications (TOSCA), which encourages modularity of applications. This is encoded in the `blueprint' files, in which the orchestrator packages the specifica-tions of the user parameters, the input and output files, as well as the binary files to be transferred into the computational re-sources. A fragment of a blueprint is shown in Listing 1. In particular, Cloudify provides a user-friendly GUI that allows to easily manage blueprint files and associat-ed deployments (Figure 6). Moreover, the blueprint file allows the orchestrator to delegate the required staging of input and output data to the Data Mover component.

For instance, the *data_mover_options* fields in the blueprint specify the files to be transferred, as well as the source, desti-nation, and the user credentials.

In addition, the orchestrator's API allows the execution to be requested from a user-friendly web interface as shown the next sections.



Source: Screenshot from http://cloudify-api.test.euxdat.eu/console

Figure 6: Cloudify GUI. Cloudify offers an intuitive GUI that enables users to upload blueprints, create deployment. The GUI also gives an overview of the number of compute nodes and running execution.

Hybrid orchestrator

Many applications that run on HPC are usually part of bigger workflows that run in the cloud, such as those with tightly-coupled data or extensive big data analytics. In a similar fashion as for micro-services applications running solely in the Cloud, an automatic hybrid deployment and management of a modularized application in HPC and Cloud is expected to optimize the overall performance and enable new development architectures. In order to address this technical gap, the AI, Data & Robotics Unit (former *Advanced Parallel Computing lab*) Lab at ATOS Research and Innovation (ARI) Spain has recently developed within the EUXDAT project a plugin for Cloudify called Crou-pier (Carnero and Nieto, 2018) written in Python that adapts the Cloudify orchestrator algorithm for the management of HPC resources, so that Cloudify can orchestrate a hybrid environment including both Cloud and HPC. Croupier has been developed to essentially bring the latest and greatest features that have been enabled by Cloud architectures, such as modularity, interoperability, software as a service (SaaS), infrastructure as code (IaC), con-tinuous integration and deployment (CI/CD), to the world of HPC. Thanks to Croupier, it is for example possible to run batch applications on both HPC and Cloud.

Data mover

The size of the data files requires an efficient transfer method for transferring them into the computational resources.The current network protocol used for such pur-poses in centers like HLRS is GridFTP. Published results show that GridFTP pro-vides better performance, between 5 and 10 times faster, than the standard FTP protocol (Esposito et al., 2003). Other ad-vantages of using GridFTP include the security provided based on x509 certificates and the capability to carry out third party transfers (Figure 7).

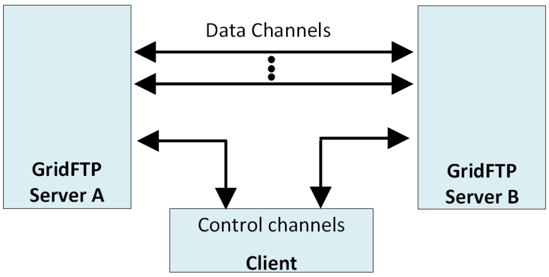
```
node_templates:
job:
    type: croupier.nodes.job
    job_options:
        type: "SRUN"
        command:"coordinates.txt"
        nodes: 100
        max_time: "04:00:00"
data_mover_options:
    workspace: wsdata
    create_ws: TRUE
    source: "ATOS"
    destination: "HLRS"
    source_input: demo_cloud_folder
    dest_output: demo_hpc_folder
    grid_userkey:
        --BEGIN ENCRYPTED PRIVATE KEY--
        ...
        --END ENCRYPTED PRIVATE KEY--
    grid_usercert:
        -----BEGIN CERTIFICATE-----
        ...
        -----END CERTIFICATE-----
```

Source: own research and processing

Figure 7: Example of a fragment of a TOSCA blueprint file.

Figure 8 shows the main elements and their communication in a GrdiFTP third party transfer. The concept consists of a client point from where the transfer is re-quested, but the data does not have to traverse it as in other protocols. Instead, the data is directly transferred between servers (servers A and B in the figure) using multiple communication channels in parallel. This protocol already proved better performance on data transfers performed over the internet because it allows using the high bandwidth available in the communication channels in the servers, by multiplexing the traffic over multiple channels in the internet that have lower bandwidth. It is highly desired to avoid using any intermediate storage, because it requires storage space, which slows down the data transfer. For that reason, we ex-plored the state of the art on large-scale data management solutions that support the GridFTP network protocol.

The most advanced tool for using GridFTP that we could find was Rucio. Rucio is an open source tool developed within the context of the ATLAS project for manag-ing big data at the European Organization for Nuclear Research (CERN). It is cur-rently used to transfer more than one petabyteworth of data per day, and more than one million files per day (Serfon, C. et al., 2019). However, after performing detailed evaluation we identified two ma-jor issues in using Rucio and therefore decided to develop our own plugin instead. The first issue is that Rucio does not allow running third party transfers. The second one is that Rucio cannot handle recursive directory transfers, when defining the des-tination folder (non-deterministic model). Although an alternative exists (determinis-tic model, where Rucio creates folders labeled with an alphanumeric string), it still makes it difficult to integrate the tool with our framework.



Source: own figure based on the GridFTP protocol (www.globus.org)

Figure 8: GridFTP direct third party transfer. Client requests a site-to-site direct data transfer. The transfer is multiplexed among multiple channels.

The development of the Data Mover plugin in place of usring Rucio allowed us to fulfill our project requirements as well as its integration with the other compo-nents. Its implementation is based on the use of the Globus-Client (SURFsara, 2015) and uberFTP (NCSA, 2020). The Data Mover plugin has been developed within the EUXDAT project and is now an inte-gral part of the Croupier plugin for Cloudi-fy (Montañana and Gorroñogoitia, 2020a). The Data Mover plugin supports GridFTP direct data transfers between data sources and computational resources without intermediate staging, as required for the previously described pilot and use cases.

HPC and Cloud Computing

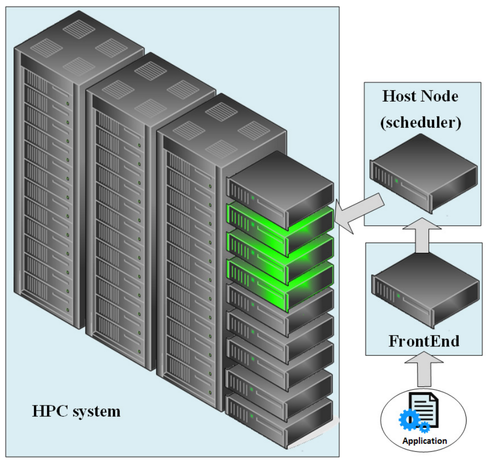
HPC and Cloud computing systems exhibit differences in terms of performance and cost. Moreover, the implementation of a given application on either of these sys-tems may differ significantly in order to achieve the best performance and resource utilization.

An HPC system consists of a large number of compute nodes that are physically close to each other (ideally all in the same room) with a high-performance interconnection network running between them. Figure 9 outlines the different nodes a user application goes through when it is submitted to an HPC system.

HPC systems rely on low-latency communication to share computational results between compute nodes. Moreover, HPC systems tend to have higher costs than Cloud systems for executing applications, due to the high cost of the interconnection network as well as the maintenance cost incurred due to having to run cooling sys-tems needed to dissipate the large heat generated by the hardware that is physically located in a relatively small space.

On the other hand, Cloud computing systems consist of different types of physical hardware (e.g., networking equipment, load balancers, servers) that can be located in different geographical locations. Virtual-ization is also typically employed in such systems in order to connect servers together, and also to divide and abstract resources in order to make them accessible to users. Cloud computing systems typically do not require a high-performance interconnection network between compute nodes, and as these are not located in the same space Cloud computing systems do not incur additional costs for cooling systems.

Applications executed in HPC centers typically show better performance than those executed in the Cloud. However, in order to make efficient use of resources additional effort is needed in order to `parallelize' applications. Parallelizing an application essentially entails spreading the workload among different computing nodes. On the other hand, applications to be executed in the Cloud can be easier to implement as it is possible to simply submit replicas of the same application as different independent jobs, each targeting a different portion of input data.



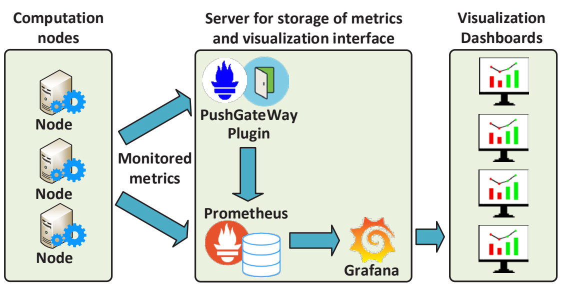
Source: own research and processing

Figure 9: Different nodes involved in the execution of an application in an HPC system. User uploads the application on a `frontend' node. The job is then dispatched to a scheduler node that manages where and when it will be run. The computation nodes, highlighted in green, are responsible for running jobs.

Although using HPC systems may lead to greater performance than Cloud computing systems, it requires higher implementation efforts and costs for execution. Hence, there is not a general agreement on which system to run a given application on, and it is up to the user to decide what to trade between implementation effort and execution cost.

Monitoring

In order to improve future application executions, the metrics related to the utilization of different resources are registered into a monitoring server such as Prometheus. Using an open source system moni-toring and alerting toolkit such as Prome-theus facilitates the decision of where to allocate future task requests depending on specific user constraints, such as reduced computation time or cost. Figure 10 shows the overall monitoring process, which includes collection of metrics at the computation nodes, storing the metrics in a Prometheus server (with metrics being pushed by the Pushgateway plugin to the server if they cannot be directly sampled at the computation nodes by Prometheus), and accessing the stored data through the Grafana visualization interface. This system allows to inspect metrics in real time as soon as they are stored in Prometheus.



Source: own research and processing. Logos of Pushgateway, Prometheus, Grafano took from https://prometheus.io and https://grafana.com

Figure 10: Collection and access to the monitored metrics.

Once the computation is completed and the monitoring metrics have been collected, the results are moved into a repository that can be accessed by the user, and the user is notified.

Use cases and pilots

The three EU projects presented above are focused on the development and testing of solutions for the field of agriculture. Agriculture is a key player in economic and political stability.

Because of its importance, governments are funding the development of solutions data access systems, geoprocessing, and tools for decision making.

The different uses cases demonstrate the capacity of the HPC solutions proposed across the projects.

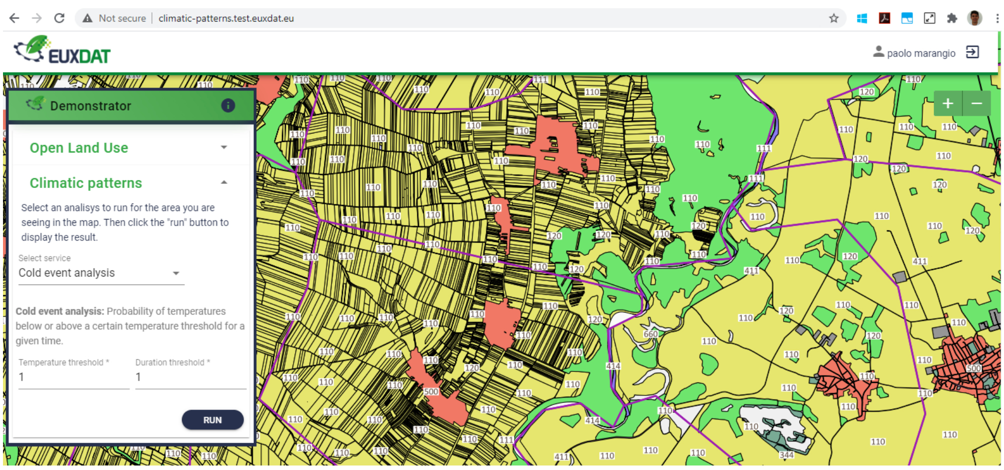
The use cases cover a wide range of real-life applications ranging from detection of weather conditions, humidity or crop dis-eases, to precision agriculture, livestock farming, and exploration. In the next sections, we provide a brief description of a selection of use cases. It should be noted that access to the implementations of the use cases is currently limited to consorti-um partners. Towards the end of the respective projects, the use cases will be made available for the end users commu-nities.

OpenLand monitoring and sustainable magement

This use case aims at developing a deep learning algorithm that uses a range of input data for predicting soil and crop status. The input data includes images generated by multirotor UAV systems with a hyperspectral camera as well as Earth observation and meteorological data.

3D Farming

This use case focuses on analytics models within the context of spatial analysis for farming (e.g., locating the highest productivity zones). It will



Source: screenshot from http://climatic-patterns.test.euxdat.eu

Figure 11: Web interface of the OpenLand monitoring pilot.

provide 3D visualization for the obtained results, which is especially helpful for understanding different soil parameters such as amount of water, and solved particles and nutrients (Figure 11).

Organic soya yield and protein content prediction

There is a strong interest in the predictive analytics of soybean farming, mainly because the EU is strongly dependent on other continents for sourcing plantbased proteins. In light of this, this use case develops methods for predicting maps indicating soybean yield and protein-content based on crowd-sourced data, satellite imagery, and additional information if available, such as electromagnetic soil scans and other sensory data.

Climate-smart predictive models for viticulture

This use case addresses the development of complex, highly-nonlinear models for vine and grape growth, which rely on a large number of variables that have been shown to affect the quality and quantity of the produced yields. The range of input data includes soil/elevation maps, earth observations, genomics, chemical analysis, environmental and climatic data.

Climate services for organic fruit production

This use case aims at helping with the prevention of damaging effects caused by frost and hail. The solution under development focuses on providing risk probability mapping calculated based on models obtained by machine learning techniques.

In order to train the predictive model, a wide range of data sources are used including but not limited to climate instability indices, digital terrain models, in-situ environmental and climatic data, and satellite images.

Optimizing computations for crop yield forecasting

This use case aims at developing a crop yield monitoring tool that can be used for agricultural monitoring (e.g., early warning and anomaly detection), index-based insurance (index estimates), and farmer advisory services. Its goal is to compute a productivity estimation based on cropping systems model and a combination of different datasets, such as ingest crop, soil, historic weather, and weather forecasts data. The computation underlying this use case becomes more challenging as the amount and resolution of available data are increased.

Evaluation

Next, we describe the pilot used to validate the framework. Results of the perfor-mance of the pilot are shown after that.

Case study: Agroclimatic-zones pilot

In this section, we focus on the *agroclimatic-zones pilot* for the validation of the framework proposed in this paper. We decided to use this pilot because at the time of writing it was one of the most mature pilots available across all the projects. From a computing point of view, the algo-rithm associated with this use case is also relatively 'simple' while still holding the potential to be parallelized using one of the parallel programming frameworks for execution on HPC described above.

Currently, the available maps of climatic zones are very generic and exhibit low granularity. Although they are able to display some differences in topography between areas, the areas shown by such maps tend to be quite large and do not include, for example, seaside buffer zones, weather divides, and South-North differences.

The idea of this algorithm is to create a classification system that allows to characterize land areas as different *agroclimatic* groups, based on long-term climate data, land cover, and topography information.

The goal of the algorithm is to generate local climate maps that take into account general weather conditions (large-scale weather models), local topography (with North/South slopes), buffer effects (such as lakes, sea, or swamps) and soil types.

The tool is primarily intended for:

- Agricultural extension counselors, or technical farm organizations wishing to make an investment in frost protection, irrigation, etc.
- Insurance and other financial institutions wishing to make decisions on quality and risk of agricultural invest-ments.
- Researchers interested in making decisions related to field trial (climatic) representativeness.
- Researchers and advisors interested in checking the impact of climate change on a given area and making decisions about future management strategies.

The expected frequency of use for the tool is once per year, while the type of data queries could be either local or regional (e.g., comparisons of several sites). By using the proposed tool it will be possible to predict long-term climate changes, which will in turn allow to make better-informed long-term decisions about crops and use resources more efficiently. One example of this is frost protection. In the past few years, significant parts of the Central EU Orchard and Vineyard industry have been affected by late frosts, which required making critical decisions about anti-frost protection measures, varietal changes, and risk mitigation strategies (Vitasse and Rebetz, 2018). In the rest of the paper, the application of the tool for computing frost-related information that may be useful in frost protection and management will be presented.

Inputs to the algorithm

The algorithm takes as inputs two different types of data as well as a set of input parameters provided by the user. The first piece of input data is meteorological data in the ERA5-Land (ECMWF/CDS) or NEMS30 (Meteoblue AG) format/model. The data is encoded in NetCDF files used as input data in the algorithm. The second piece of input data is optional and it includes topography maps (EU-DEM format) and land cover/soil maps (Joint Research Center or Open Land Use Map).

The input parameters provided by the users are:

- Area of interest encoded by polygon drawn over map presented to the user
- Start year
- End year
- Probability of first/last frost day
- Frost temperature (in degree Celsius)
- Daily hours with minimum tempera-ture: start hour, end hour (0-23)
- Length of stretch of last and first frost days to be found for a single year

Output of the algorithm

The output of the algorithm is a set of values corresponding to *agroclimatic* variables calculated based on the input data and user input parameters. For every point in the polygon specified by the user, the algorithm computes the following:

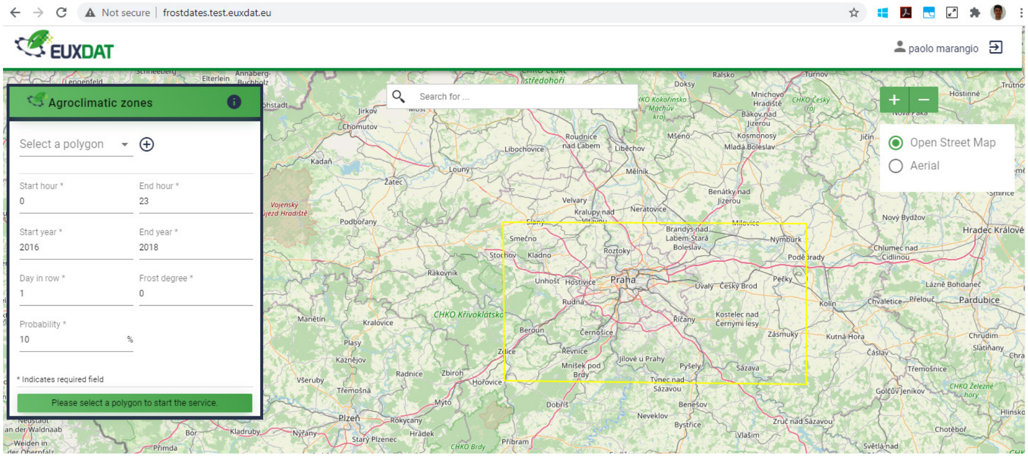
- Last spring frost date
- First fall frost date
- Length of frost-free season
- Number of frost days
- Average number of frost days for the period spanning the start and end year

The output information is stored as a GeoJSON file.

Proposed parallelization strategy

At the smallest level, the *agroclimatic-zones* algorithm calculates, for a given position on the map identified by a pair of latitude and longitude values, the first and last frost date for each year over a user-defined range of years. From a programmatic point of view (Figure 14). This is achieved through the sequential invocation of three nested functions, with information for a single year and grid point in latitude and longitude dimensional space being computed by the innermost function (*findfrostdates*), followed by an intermedi-ate function that aggregates this information across multiple years (*frostdateyearly*). Finally, the outermost function invokes the two inner functions over mul-tiple points in latitude and longitude dimensional space (*frostdatesplaces*). Since for a typical execution of the algorithm the number of years (maximum value is 37) tends to be much smaller than the total number of pairs of latitude /longitude val-ues (maximum value for this study was 70 longitude values x 27 latitude values = 1890 grid points), it seemed more intuitive to perform a parallelization of the outermost function *frostdatesplaces*, such that each MPI process would essentially be responsible for calculating the output in-formation for an independent pair of latitude/longitude values. Parallelizing this function would not break the code or cause bottlenecks because the computations for different pairs of latitude/longitude values are independent and can, therefore, be performed asynchronously. Hence, we essentially used MPI to schedule parallel execution of an identical computation (with the only difference being the input grid point of latitude/longitude values) over multiple processes, rather than using point-to-point communication for directly improving the performance of the serial/sequential *agroclimatic-zones* algorithm.

Even if the serial application was written in Python, introducing the proposed changes was easily achieved by using the Message Passing Interface (MPI) for Python package (Dalcin, 2019) (Figure 12).



Source: screenshot from http://frostdates.test.euxdat.eu

Figure 12: Web interface of the agroclimaatic zones pilot.

Results and discussion

Application requirements evaluation for execution of parallel agroclimatic-zones algorithm in HPC

The resource requirements for the *agroclimatic-zones* can be evaluated considering the infrastructure in which the algorithm will be implemented and finally executed. All of the projects have access to the Hewlett Packard Enterprise Apollo 9000 Hawk supercomputer available at the HPC Center in Stuttgart (HRLS). Table 1 shows the key features of Hawk, which was launched in February 2020.

Name	HPE Apollo 9000 Hawk
Number of node	5,632
Number of cores	720,896
Peak performance	26 Petaflops
Disk storage capacity	25 PB
Interconnection net	InfiniBand HDR (200Gbit/s)
Power consumption	2112 KW, to be increased

Source: https://www.hlrs.de/systems/hpe-apollo-9000-hawk

Table 1: Characteristics of the HLRS Hawk HPC system.

The simultaneous use of HPC systems by a large number of users requires that each user's execution request includes a specification of the number of computing nodes and software to be used. After a request for execution is submitted, it is queued until all the resources required for the execution become available. Most HPC centers need to have very high usage in order to be viable. This, however, means that the waiting time for the execution of a given application can range from a few minutes to a few days depending on the workload of the HPC center and the resource requirements specified in a given request. Obviously, the user is only billed for effective compu-tation time, not for time spent waiting in the queue. Hence, the required computation time is an important aspect to consider when executing geoprocessing applications in such infrastructures. In this respect, an option is to upload the required data to the HPC system prior to execution as this can save a significant amount of computation time (and in turn cost).

```
input:
 startlat, startlon, endlat, endlon, startyear, en-dyear,
 probability, frostdegree, starthourday, endhourday,
 dayinrow
output:
 firstfrosday, lastfrostday frostfreeperiod,
 numbfrostdays, avgnumbfrostdays
begin
 call function frostdatesplaces
   loop over lat & lon
     call function frostdatesyearly
       initialize nmbfrdayslist = empty list
       loop over years
         call function findfrostdates
           initialize numbfrostdays = 0
           initialize lastfrostday = 0
           initialize firstfrostday = 0
           loop over days between Jan and Jul
             initialize daymin = 50
             loop over hours
               calculate currentemp
               if currenttemp < daymin:
                 daymin = currenttemp
```

Source: own research and processing

Figure 13: Pseudocode of serial agroclimatic-zones algorithm (to be continued).

```
    if daymin <= frostdegree:
      numbfrostdays += 1
  output lastfrostday
  loop over days between Jul and Jan
    initialise daymin = 50
    loop over hours
      calculate currentemp
      if currenttemp < daymin:
        daymin = currenttemp
    if daymin <= frostdegree:
      numbfrostdays += 1
    output firstfrostday
    output numbfrostdays
    frostfreeperiod = (see next line)
    firstfrostday - lastfrostday
    output frostfreeperiod
    append numbfrostdays to nmbfrdayslist
  avgnumbfrostdays = mean(nmbfrdayslist)
  output avgnumbfrostdays
```

Source: own research and processing

Figure 13: Pseudocode of serial agroclimatic-zones algorithm (continuation).

Table 2 shows the preliminary requirements of two applications belonging to use cases from different projects. We chose these applications because their resource requirements are common among all of the listed use cases. In particular, the estimated size of data to be transferred and the computational load of the applications for computing *agroclimatic-zones* and land morphometry characteristics are shown. Since a use case is composed of a series of geoprocessing applications, the computational and data storage requirements of a use case presented here correspond to the accumulation of the analyzed requirements of the individual constituent applications.

While consideration must be placed in evaluating application requirements before execution, the execution time of an application is also an important aspect that factors in the decision of which infrastructure the application should ultimately be run in. For instance, if a farmer needed to know whether the next morning's temperature was going to be below 27 degrees (Muhollem, 2017), the farmer would need to receive the application output before the morning would come in order to successfully safeguard his blossoming crop. For such applications, it is more suitable to carry out the execution on HPC rather than Cloud since the computation will be completed earlier on HPC (assuming a suitable level of parallelization has been introduced).

Applications	Agroclimatic-zones frost date calculation	Morphometry characteristic calculation
Storage requirements	316 MB (ERA5-Land Czechia)	25 GB (Austria Area) 1 TB (Full Europe)
Computation time in core-hours	70 (Czechia)	3000 (Full Europe)

Source: Pavel Hájek (http://www.wirelessinfo.cz) and Dr. Karl Gutbrod (https://meteoblue.com)

Table 2: Requirements of selected applications.

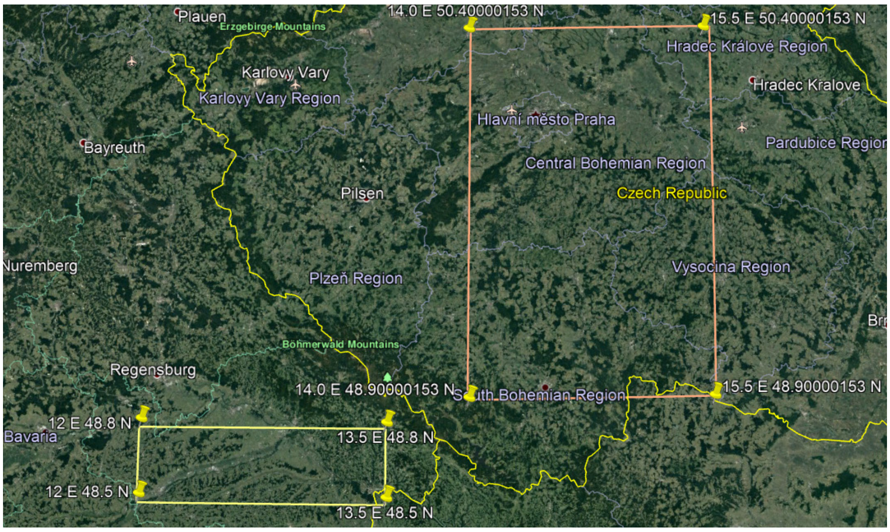
Results of benchmarking tests of parallel *agroclimatic-zones* algorithm

In order to properly benchmark the paralel *agroclimatic-zone*s algorithm, we strived to test it on an input with size and complexity consistent with those to be found when the algorithm would be in production. This required making some choices in terms of the number of years worth of data and the number of pairs of latitude/longitude values to be processed.

After some experimentation, it was found that the serial agroclimatic-zones algo-rithm was able to process 12-15 years worth of input data and 64-256 pairs of latitude/longitude values in the order of 2-85 hours. This range of execution time and problem size seemed like a reasonable starting point for experimentation, while also allowing us to more clearly showcase the capability of HPC to improve perfor-mance of a typical medium to large scale application. Therefore, all benchmarking experiments presented here were conducted with either 64 (i.e., smaller input problem size) or 256 (i.e., larger input problem size) pairs of latitude/longitude values (Figure 14) and 12 years worth of input data.

Benchmarking of the parallel algorithm was performed in HLRS Hawk. A standard compute node from this infrastructure con-sists of a single 2.25GHz, 64-core AMD Epyc Rome 7742 processors with 256 GB memory. Each of the cores in the proces-sor supports 2 hardware threads (also known as hyperthreading), meaning that a single node can execute up to 128 threads.

In order to determine the performance of the parallel algorithm, the total execution time of the program (after all MPI processes had finished execution) was measured. The parallel algorithm was run on 1, 2, 4, 8, 16, 32, 64, 128, 256 and 512 MPI processes on one node with the exception of 256 and 512 processes which were run on 2 an 4 nodes. As shown in Figure 14, the parallel algorithm ran with 64 pairs on input data reached the lowest execution time of approximately 2min 30 sec



Source: Screenshot from http://climatic-patterns.test.euxdat.eu

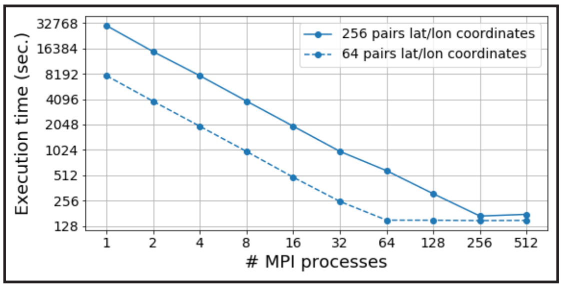
Figure 14: Map showing boundaries of grid of latitude and longitude values processed by the *agroclimatic-zones* algorithm representing a relatively small (shown by yellow rectangle) and large (shown by orange rectable) problem size in this study. Assuming a spatial resolution of 0.1x0.1° between these boundary values, the yellow rectangle corresponds to 64 pairs of latitude/longitude values (or grid points), while the orange rectable correspondes to 256 pairs of latitude/longitude values.

with 64 MPI processes. When run with a input problem size of 256 pairs of lati-tude/longitude values, the parallel algorithm managed to record the same low execution time as with the smaller input problem size, but using 256 MPI processes. Considering that the serial algorithm reported an execution time of approxi-mately 2h10min with 64 input pairs of latitude/longitude values, and 8h with 256 pairs, the parallel algorithm achieved a maximum speedup of approximately 52 times with the smaller input problem size, and speedup of approximately 182 times with the larger input problem size.

Overview of metrics collected during program execution in HLRS Hawk

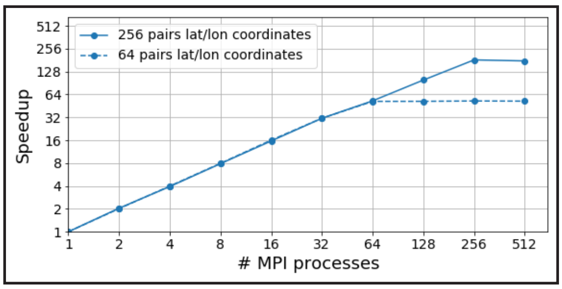
The metrics collected in Prometheus in the current implementation of the parallel *agroclimatic-zones* algorithm are the amount of data transferred by the Data Mover, and the average bandwidth on each request of transfer of a set of files (i.e., it can consist of multiple files or a single one). Additional metrics related to the performance of the applications are also automatically collected.

The Data Mover was tested by transferring test files with different sizes (i.e., 100MB, 1GB, and 10 GB) between the Cloud resources of ATOS in France and HPC resources of HLRS in Germany. The test files contained random data in order to avoid data compresion. The typical band-width measured with GridFTP transfers was between 70 and 90 MB/s. However, it is not possible to use these results to pre-dict the future performance of a given transfer because it will ultimately depend on the network load in the internet as well as in the data centers (Figure 15 and 16).



Source: own research and processing

Figure 15: Number of MPI processes vs. Execution time. Input data is 64 (shown by dashed line)or 256 (shown by the solid line) pairs of latitude/longitude values and 12 years worth of ERA5-Land data.



Source: own research and processing

Figure 16: Number of MPI processes vs. Speedup. Input data is 64 (shown by dashed line) or 256 (shown by the solid line) pairs of latitude/longitude values and 12 years worth of ERA5-Land data.

Analysis of results

The infrastructure platform proposed within the context of the EU projects discussed in the paper currently satisfies all the re-source requirements for the agroclimatic-zones pilot, and no deficiencies could be detected.

The benchmarking tests of the parallel algorithm used for deploying the *agroclimatic-zones* pilot have clearly demonstrated the power of HPC for parallelizing data-intensive applications. Thanks to the proposed parallelization strategy and by running the application on multiple compute nodes of an HPC cluster, the final parallel algorithm was 52 times faster than the original sequential program for a relatively small yet realistic input size. Put in a different perspective, these results al-ready means that a user can now execute his/her program in a couple of minutes instead of several hours. Considering that an even greater speedup (i.e., approximately 182 times faster than sequential pro-gram) was achieved when problem size was quadrupled (i.e., 256 pairs of latitude/ longitude values is doubled) as well as the number of MPI processes, these results indicate that the parallel algorithm scales well with problem size. This strong-ly suggests that the newly developed algo-rithm presented in this paper is a very efficient, parallel algorithm that should be of general interest to the geoprocessing community.

Based on the results and the evaluation presented in this study it can be argued that the proposed framework simplifies the deployment and execution of geoprocessing tasks. Thanks to its data moving approach and the use of HPC resources, the framework is able to achieve an efficient transfer of data and computation in a significantly smaller amount of time, therefore also reducing costs. Based on the current body of knowledge the proposed framework seems to be very cost-effective for geoprocessing and is particularly suitable for large projects such as large scale studies conducted by governments. It is also attractive for companies interested in selling the results of geopro-cessing to small customers that do not have access to the data or the software necessary for running applications by themselves.

Conclusions

In this paper, an open source framework underpinned by an infrastructure suitable for HPC and Cloud computing of geopro-cessing services has been described. We have demonstrated that the infrastructure can support the execution of realistic use cases within the context of several EU projects, and achieve large speedup (up to 182 times) when running data-intensive applications.

The solutions being developed by the EU projectsshowcased in the paper will greatly support improving farming performance and competitiveness. This is not only because the developed tools are fit for purpose, but also because they leverage time-efficient computational resources. These tools will exhibit a simplified access for non-technical users. They are attractive also for customers that do not have access to the data, software or hardware needed. Moreover, the intention is that the developed platforms will stay operational after the end of the respective projects. In particular, the partners in the projects are in-terested in using them for selling their products, such as datasets and weather forecasting services directly to farmers after the respective projects are over. In order to ensure this, the consortium part-ners are committed to perform the roles of software, HPC and Cloud platform provid-ers after the projects are over.

Additionally, it should be noted that the developed platform for agriculture geo-processing is also suitable for other pur-poses than agriculture, such as providing optimum paths through transportation networks, predicting disasters like wildfire and flooding, or the effects of a storm. Considering this broader scope, potential users can therefore also include local au-thorities interested in urban and regional planning and water management, or insurance companies interested in risk prevention or disaster resilience.

Acknowledgments

This work has been carried out within the context of the following projects: *European e-infrastructure for extreme data ana-lytics in sustainable development* (EUXDAT); *Fostering precision agricul-ture and livestock farming through secure access to large-scale HPC-enabled virtual industrial experimentation environment empowering scalable big data analytic* (CYBELE); *Open interoperable platform for unified access and analysis of Earth observation data* (EOPEN).

Further information about the projects is available at the respective web pages (Nieto et al., 2020; Vingione et al., 2020; Davy et al., 2020). The research leading to these results has received funding from the European Unions Horizon 2020 Research and Innovation Programme, grant agreements n. 777549, 825355, 776019, respectively.

We wish to thank Dimitrij Kozuch (Plan4all), Pavel Hájek (WirelessInfo), Jiří Valeš (University of West Bohemia) and Dr. Karl Gutbrod (Meteoblue AG) for providing the applications described or showcased in this paper, as well as the estimation of their requirements. We wish to thank Naweiluo Zhou for providing an up to date description of the CYBELE project. We would also like to thank the (30+) research institutes and enterprises across several EU countries that have or are still collaborating on these three Euro-pean projects for their valuable work.

Corresponding authors
José Miguel Montañana Aliaga, Ph.D, Senior Researcher
Höchstleistungsrechenzentrum Stuttgart (HLRS)
Nobelstraße 19, 70569 Stuttgart, Germany
E-mail: jmmontanana@gmail.com

References

[1] Amdahl, G. M. (1967) "Validity of the single processor approach to achieving large scale computing capabilities", In *Proceedings of the Spring Joint Computer Conference (NY, USA), AFIPS '67 (Spring)*, Association for Computing Machinery, p. 483-485. DOI 10.1145/1465482.1465560.

[2] Carnero, J. and Nieto, F. J. (2018) "Running simulations in HPC and cloud resources by implementing enhanced Tosca workflows", In *2018 International Conference on High Performance Computing & Simulation (HPCS)*, pp. 431-438. DOI 10.1109/HPCS.2018.00075.

[3] Dalcin, L. (2019) "*MPI for Python*". [Online]. Available: https://mpi4py.readthedocs.io/en/stable/index.html [Accessed: 15 Sept. 2020].

[4] Davy, S. (2020) "*CYBELE Fostering Precision Agriculture And Livestock Farming Through Secure Access To Large-Scale Hpc-Enabled Virtual Industrial Experimentation Environment Empowering Scalable Big Data Analytic*". [Online]. Available: https://www.cybele-project.eu [Accessed: 15 Sept. 2020].

[5] Esposito, R., Mastroserio, P., Tortone, G. and Taurino, F. M. (2003) "Standard FTP and GridFTP protocols for international data transfer in Pamela Satellite Space Experiment. In *Proceedings from the 13th International Conference on Computing in High-Enery and Nuclear Physics (CHEP 2003).*

[6] Li, Z. (2020) "Geospatial Big Data Handling with High Performance Computing: Current Approaches and Future Directions", In *High Performance Computing for Geospatial Applications*, pp. 53-76. E-ISBN 978-3-030-47998-5, ISBN 978-3-030-47997-8. DOI 10.1007/978-3-030-47998-5_4.

[7] Mineter, M. J., Dowers, S. and Gittings, B. M. (2000) "Towards a HPC Framework for Integrated Processing of Geographical Data: Encapsulating the Complexity of Parallel Algorithms", *Transactions in GIS,* Vol. 4, No. 3, pp. 245-262. E-ISSN 1467-9671. DOI 10.1111/1467-9671.00052.

[8] Montañana, J. M. (2010) "Providing Fault Tolerance in Interconnection Networks for PC Clusters: Efficient Mechanisms", Lap Lambert Academic Publishing. ISBN-13 978-3838318905, ISBN-10 9783838318905.

[9] Montañana, J. M. and Gorroñogoitia, J. (2020a) "*Data mover plugin provides support for GridFTP data transfers to croupier cloudify orchestrator*". [Online]. Available: https://github.com/ari-apc-lab/croupier/tree/euxdat/croupier_plugin/data_mover [Accessed: 20 Sept. 2020].

[10] Montañana, J. M., Hervás, A. and Hoppe, D. (2020b) "HPC-Enabled Geoprocessing Services Cases: EUXDAT, EOPEN, and CYBELE European Frameworks", In *Proccedings of the 12th International Conference on Advanced Geographic Information Systems, Applications, and Services (GEOProcessing)*, pp 31-35.

[11] Muhollem, J. (2017) "*Warm winter has put state's apple crop at risk, expert warns", Pennsylvania State University*. [Online]. Available: https://phys.org/news/2017-03-winter-state-apple-crop-expert.html [Accessed: 20 Sept. 2020].

[12] NCSA (2020) "*uberftp - GridFTP-enabled client*". Linux man page. [Online]. Available: https://linux.die.net/man/1/uberftp [Accessed: 20 Sept. 2020].

[13] EUXDAT (2020) "*EUXDAT European e-Infrastructure for Extreme Data Analytics in Sustainable Development*". [Online]. Available: https://www.euxdat.eu [Accessed: 20 Sept. 2020].

[14] Perakis, K., Lampathaki, F., Nikas, K., Georgiou, Y., Marko, O. and Maselyne, J. (2020) "CYBELE – Fostering precision agriculture & livestock farming through secure access to large-scale HPC enabled virtual industrial experimentation environments fostering scalable big data analytics", *Computer Networks*, Vol. 168, ISSN 1389-1286. DOI 10.1016/j.comnet.2019.107035.

[15] PESSL INSTRUMENTS GMBH (2020) "*Stations and datalogger*". [Online]. Available: http://metos.at/micrometos-clima [Accessed: 15 Aug. 2020].

[16] Serfon, C., Barisits, M., Beermann, T., Garonne, V., Goossens, L., Lassnig, M., Nairz, A. and Vigne, R., ATLAS Collaboration (2019) "Rucio, the next-generation Data Management system in ATLAS", *Nuclear and Particle Physics Proceedings*, Vol. 273-275, pp. 969-975. ISSN 2405-6014. DOI 10.1016/j.nuclphysbps.2015.09.151.

[17] SURFsara (2015) "*Globus client. Grid Documentation v1.0.*" [Online]. Available: http://doc.grid.surfsara.nl/en/latest/Pages/Advanced/storage_clients/globus.html [Accessed: 15 Aug. 2020].

[18] EOPEN (2020) "*EOPEN Open interoperable platform for unified access and analysis of earth observation data*". [Online]. Available: https://eopen-project.eu [Accessed: 20 Sept. 2020].

[19] Vitasse, Y. and Rebetez, M. (2018) "Unprecedented risk of spring frost damage in Switzerland and Germany in 2017", *Climatic Change*, Vol. 149, pp. 233-246. E-ISSN 1573-1480, ISSN 0165-0009. DOI 10.1007/s10584-018-2234-y.

[20] Zhang, J. (2010) "Towards personal high-performance geospatial computing (hpc-g): perspectives and a case study", In *Proceedings of the ACM SIGSPATIAL International Workshop on High Performance and Distributed Geographic Information Systems (HPDGIS)*, pp. 3-10. DOI 10.1145/1869692.1869694.

Digitalisation in the Food Industry – Case Studies on the Effects of IT and Technological Development on Companies

Judit Nagy, Zsófia Jámbor, Anna Freund

Corvinus University of Budapest, Hungary

Abstract

In recent decades as a result of the development of information and communication technologies (ICT) and the Internet, we have witnessed major changes in companies. The ICT support of the processes is becoming more and more extensive and comprehensive, which enables the realization of digitalisation. The interconnection of processes, machines, people in a single network makes another level of optimisation available. The changes turned up by digitalisation are not only technological, but they also have an impact on the company's organisation and strategy. Our study aims to create an analytical framework and map the opportunities that digitalisation promises in the food industry and the organisational changes that ICT and technological development bring, with special emphasis on the impact on strategy, employees, and corporate culture. Our results show that companies are not consciously engaged in digitalisation yet, but they exploit their opportunities and make improvements in this sense. Adaptation of digital solutions is often forced by the labour shortage, the pressure to achieve higher efficiency and thus to remain competitive and to service the growth strategy.

Keywords

Digitalisation, ICT development, food industry, strategy, interviews.

Nagy, J., Jámbor, Z. and Freund, A. (2020) "Digitalisation in the Food Industry – Case Studies on the Effects of IT and Technological Development on Companies", *AGRIS on-line Papers in Economics and Informatics*, Vol. 12, No. 4, pp. 77-92. ISSN 1804-1930. DOI 10.7160/aol.2020.120406.

Introduction

The development of information and communication technologies (ICT), database systems, and increasing digitalisation is a hot topic not only in the corporate but also in the academic sphere. Robots and automation have long existed, but the Internet (or any internal company network like this) is revolutionizing process management by connecting them into a network. Thanks to the increasing digitalisation, the company's devices, and machines, and in some cases the products themselves can connect and cooperate, realizing Industry 4.0 (I4.0). Devices such as sensors, RFID chips, 3D scanners, cameras, and robots generate data that literature calls big data. It is a huge challenge to store, process, and interpret, and is essential for the implementation of I4.0 (Hermann et al., 2016; Ilie-Zudor et al., 2011). Digitalisation is not only an ICT development, its extensive use in the company processes also affects the organisation (Horváth and Szabó, 2019), employees (Frey and Osborne, 2017), and business partnerships (Pagani and Pardo, 2017).

Most of the published studies deal with digitalisation from a technological point of view (Dworschak and Zasier, 2014; Hermann et al., 2016) or a theoretical point of view (Fettig et al., 2018; Dalenogare et al., 2018) however, the number of studies presenting specific company cases is limited. The impact of digitalisation and I4.0 is most significant in the electronics and machine industry, most of the research reports illustrate their solutions (Demeter et al., 2020; Nagy, 2018; Horváth and Szabó, 2019; Gauger et al., 2017). In our study, we focus on another manufacturing sector, the food industry, which is not a typical subject of digitalisation studies. Robotics and automation have long been present in many subsectors of the food industry. Real-time data streaming, big data, and information sharing, first and foremost within the company and then at the supply chain level, are of paramount importance for food safety and tracking purposes, and industrial digitalisation solutions offer many opportunities to improve these areas.

1. World trends in the food industry

Due to global trends, it is very important to focus on the agriculture and food industry: on the one hand, to supply a growing world population, more efficient production and processing are needed, and on the other hand, guaranteeing food safety is already a global requirement. The world population has tripled since the 1950s, reaching over 7.8 billion in February 2020. This figure, although slowing, continues to grow and is estimated to exceed well over nine billion by 2050 (Chaime, 2020). The World Economic Forum estimates that food demand will increase by nearly 70% so far (World Economic Forum, 2018). Another important aspect is that with the development of the countries, due to the increase of the living standard of the population, the amount of food to be consumed changes significantly, the energy demand of the population is much higher and the daily intake of animal protein is increasing (Sen et al., 2017). Meeting growing and changing demand and reducing and relieving environmental impacts mean major challenges for farms as well as actors in both agriculture and the food industry (Demartini et al., 2018).

One of the industries most exposed to the variability of consumer demand is the food industry. Even if the quantity needed does not change dramatically, the type of products demanded - e.g. increasing trend of healthy food– causes significant changes (Carpenter and Wyman, 2017). The production of healthy foods requires new, high-quality ingredients and production methods, which can reduce sugar and/or fat content, use alternative substitutes, and retain higher levels of vitamins and minerals. The food industry should take into account for example Europe's aging societies and changing nutrition needs, and childhood obesity, as a growing problem in the developed world when developing (new) products (Santeramo et al., 2018).

In addition to consumer trends, the food industry also has to cope with pressure from retailers, which means lower prices, higher quality, a constantly renewed range of products, and, of course, unquestionable food security (Kittipanya-ngam and Tan, 2020). To address these challenges from both sides, the development of ICT, digitalisation, and I4.0 tools can provide solutions.

Regarding the structure of the paper, after we give a brief description of the world food industry trends and the possibilities of ICT-based digitalisation and I4.0, in the methodological section we present the theoretical framework we used for the analysis and the method and process of the case studies. The fourth section summarises the experiences of the company interviews, after which we conclude. Finally, research limitations and future research directions will be presented.

2. Industry 4.0 and digitalisation in the food industry

IT-based automation and robotization have long been present in many subsectors of agriculture and the food industry (e.g. pasta, dairy), while in others are only partially applicable, because of the high demand for human workforce since processes are poorly standardized or too sophisticated (meat industry, bakery products). As Kittipanya-ngam and Tan (2020) states, there is not a single ICT solution for digitalising the food industry and agriculture processes. According to Simutech (2016), machine downtime in food processing can cost up to $ 30,000 per hour, so the use of in-process predictive maintenance sensors to prevent machine failure can quickly pay off. Processes made transparent with sensors can help increasing energy efficiency, reducing rejected deliveries and waste. Agility, rapid machine change-over, and the ability to produce smaller series are important considerations when selecting new technology (Carpenter and Wyman, 2017).

Meeting the changing customer demand also transforms and accelerates the product development process. With 3D printing, the prototype of either the product itself or the advanced packaging is shortened, so consumer testing can be accomplished in a much shorter time.

According to experts (Bibi et al., 2017; Carpenter and Wyman, 2017; Bottani and Rizzi, 2008), food safety and traceability are definitely the areas where information and database systems as the representatives of I4.0 can greatly support the management of agriculture and food industry. For example, identification systems based on the EPC Global system can trace raw materials incorporated into food and/or packaging from the place of origin to the place of use. Thus, in case of any problem or product recall, the affected product series can be clearly identified and quickly withdrawn from the market (Bibi et al., 2017; Carpenter and Wyman, 2017).

Researchers see blockchain technology as an appropriate tool for food monitoring (Tian, 2016; Tse et al., 2017). As discussed in the previous paragraph, ICT based solutions, such as RFID, barcode, and wireless sensors are well-suited to the food supply chain for transparency and traceability reasons, and especially for data collection and transmission. However, we need a medium that makes this information visible (Carlozo, 2017). Blockchain raises the level of trust by making the flow of data, goods, or money transparent and traceable.

Thus, both the agriculture and the food industry can take advantage of the ICT based I4.0 applications and the solutions offered by digitalisation, in many ways. However, some technologies and information and database connected methodologies are either expensive, immature, or have little practical experience, so there is no benchmark for learning.

Our research question is how do food industry companies perceive the realization of digitalisation and ICT based improvements in their operations and organisation? The novelty of our paper is twofold. On one hand, we create a research framework, which helps systematically analyse a company's digital situation. Besides the applied ICT solutions, we invoke additional research aspects and propose to analyse digitalisation's effects on strategy, organisation, employees, and company culture. On the other hand, until now only a few academic papers have analysed the presence and effects of I4.0 and digitalisation in the food industry. Our results widen the knowledge on digitalisation in the food industry, point out its effects not only on technology but on the entire organisation, and could be benchmarks for practitioners, too.

Materials and methods

To systematically examine the impact of I4.0 and digitalisation on the food industry companies, we looked for relevant analytical aspects in the literature. The purpose of this analysis is to get to know how I4.0 and digitalisation transform the entire company as well as to call the attention of business professionals that I4.0 not only affects the technology in a company, but also the organisation itself.

The paper presents case studies that are based on four interviews and were conducted in the three largest subsectors of the Hungarian food industry. According to Yin (2011), a case study is an appropriate methodology when researchers want to understand how digitalisation affects companies and aims to gain a better understanding of what and why firms decide and do to move towards I4.0. The case studies are not intended to produce generalizable or representative results (Denzin and Lincoln, 2011), however, they might help to identify good practices in the sector and further directions for improvement.

1. Analytical framework

To get a broad picture of how I4.0 affects the organisation we tried to identify relevant analytical dimensions. We found that I4.0 maturity models involve a great variety of aspects, which seemed to be useful for our analysis, too. Many authors and consulting companies have developed maturity models that evaluate companies in several aspects before determining their degree of maturity and many of these aspects are common. We do not aim to examine the maturity of the analysed companies, because the sample size is not large enough, but some maturity dimensions can serve as a basis for our analysis because they examine the individual organisations (Table 1). In the following, three highly cited maturity models are presented and compared to determine our analytical framework.

The most cited maturity model comes from Schumacher et al. (2016), who examine companies' digital status from nine perspectives. The analytical dimensions include strategy, leadership, customers, products, operations, technology, culture, people, and governance. The study does not detail the nine assessment factors but reveals that they are built of another 62 dimensions. As we can see, Schumacher et al. analyse the effects of I4.0 on the organisation from various points of view, which suggests us not to narrow our investigation only to technology. From amongst the aspects, organisational adaptation is clearly lacking.

Geissbauer et al. (2016) have formulated a maturity model that involves cross-company digital connectivity with customers, suppliers, and other stakeholders, who are all integrated horizontally and vertically into a digital ecosystem. The scope of the analysis is very broad, besides the digital business model (change of strategy), agile IT infrastructure, and data analysis, many new aspects are emerging, including organisation, taxation, and data analysis. In their approach, I4.0 also has a comprehensive effect on organisations, besides technological improvement.

Schuh et al. (2017) developed a complex index to measure the maturity of companies in I4.0. They examine the structural characteristics of firms such as resources, organisational structure, information

Authors	*Aspects of maturity analysis*	*No. of citations (Google Scholar)*
Schumacher et al., 2018	Strategy, Leadership, Customers, Products, Operations, Technology, Culture, People, Governance	415
Geissbauer et al., 2016	Digital business models and customer access, Digitization of product and service offerings, Digitization and integration of vertical and horizontal value chains, Data & Analytics as core capability, Agile IT architecture, Compliance, security, legal and tax, Organization, employees and digital culture.	198
Schuh et al., 2017	Structure: Resources, Organizational structure, Information system, Culture Processes: Functional areas (Development, Production, Logistics, Services, Marketing & Sales)	76

Source: own edition

Table 1: Maturity models examined.

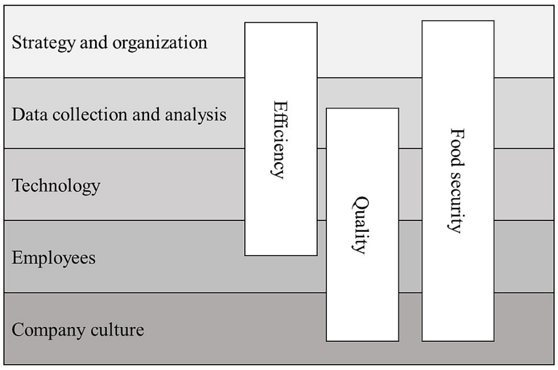
systems, and culture. Corporate processes are also reviewed along with functional areas, and performance is evaluated also along with the functional areas. The specialty of this approach is that the maturity of the functional areas underpins the maturity of the entire company. On the other hand, this approach is not fortunate, because the primary area of I4.0 development is usually manufacturing (Hofmann and Rüsch, 2017; Brettel et al., 2014), and other functional areas are usually followers if developed at all. Overall, this model can give us a complete picture of the development of a company in I4.0, since a company should not be judged solely on the level of development of its manufacturing process.

The Schuh-model combines analytical aspects that are already present in earlier maturity models in many ways. The resource category includes all movable resources - machines, products, tools, materials - including people and their abilities. The information systems category includes all formal communication solutions, including ICT (information and communication technology). The organisational structure examines the existence of an agile organisation and inter-company relationships that allow developing technology and continuous adaptation. Culture captures the soft side of all this, such as knowledge management, decision-making, corporate values, and the support of innovation. Thus, the company and the functional units are analysed according to these aspects, in the form of a questionnaire.

We do not include it in the analysis, but there are other (less-cited) maturity models that try to extend the above ones or similar to them. In their model, Lichtblau et al. (2015) propose six criteria for determining the maturity of companies. These six aspects are broken down into a further 18 in detail, very similar to Geissbauer's. According to Gracel and Łebkowski (2018), the Manufacturing Technology Maturity Model (MTMM) wants to categorize companies along with the digital support of manufacturing, but in fact, they did an extensive analysis based on eight dimensions, already mentioned in the previous models and primarily focused on manufacturing. The maturity model presented by Gill and Van Boskirk (2016) is short and concise, analysing companies from four perspectives: culture, technology, organisation, and insights. The latter involves data analysis, the ability of the company to analyse and use data from customers or processes to support the strategy.

From the maturity models presented and the underlying analytical considerations, we will use the dimensions that evenly appear in the models. Technological improvements to gain a better data collection and analysis are the basis of the developments within a company, and since the maturity models emphasize the issue, we included it into our analysis, too. The extensive technological advances – not only in data-issues but in manufacturing and other areas, too - are another characteristic projection of I4.0, which we obviously need to analyse. After these two fundamental changes that typically occur in every company, we wanted to examine the effects of I4.0 on the organisation from multiple aspects. First, we choose employees, since they are the ones who use the new technology and provide or analyse data, make decisions. Because of the probable changes in the decision-making process, we also keep it important to get to know-how strategy and organisational structure as well as the company culture changed. These aspects can also be derived from the cited maturity models.

The purpose of this paper is not to determine the degree of maturity, but to provide a structured presentation of the technological and organisational development that is taking place in the studied companies as a result of digitalisation. We will apply analytical dimensions of strategy and organisation, data analysis, technology, employees, and corporate culture to highlight how the Hungarian food industry companies move towards I4.0. The relationship between these and the most important competitive factors for the food industry is illustrated in Figure 1.



Source: own construction

Figure 1: Relationship between research aspects.

The figure summarises the five analytical dimensions along which the interviewed companies will be analysed. The three most important competitive factors – efficiency, quality, and food security – were derived from the literature review and are the largest challenges of today's food industry. We are intended to analyse how these competitive factors are affected by the changes indicated by digitalisation. The most important factors in increasing efficiency – which, according to Kürthy et al. (2016), is a serious lag in the Hungarian food industry – are the strategy which incorporates the understanding of the firm's actual situation and the adaptation, the technology which is changed as a result of educating employees, and uses the resulting data for further optimisation. Technology also has a great impact on quality, and by analysing the data, it can be further developed, faster prototype production and testing can be achieved. Commitment to quality must also be part of the corporate culture, which is reflected in the work of the employees. Digital developments in all areas serve a high level of food safety and traceability. Overall, the analytical aspects of the interviews will therefore help better understanding the three competitive factors.

2. Case studies

Following the above-identified analytical aspects, we edited interview questions to get insight into the companies' I4.0 practice, motives, and achievements. Data were collected from four food companies in Hungary, whose names will not be disclosed upon request. The most important data on the investigated companies are presented in Table 2.

The interview questions started with a strategy. We asked the companies about their market, the changes they experience, and the possible answers they are giving to these changes. he data collection, analysis, and other technological improvements were asked as to what kind of improvement projects have been carried out or planned at the companies. We also asked them about how and what extent the employees are involved in the development process, and how he changes affected them. We were curious about the organisational changes caused by I4.0, for example, new positions or the probable downsizing. The company culture was captured by questions like how employees adapted to changes, what steps the companies do to make the staff accept the new technologies.

We made 1.5-2 hours long structured interviews with company executives, and we talked about the company's I4.0 approach, the digitalisation solutions used, the projects, and the directions of development. As all interviewees were top executives of a given company, they were aware of both production and strategic issues and were considered relevant sources of information. The reliability of our research is enhanced

Company / Characteristics	"A"	"B"	"C"	"D"
Sub-sector	pasta manufacturing	dairy production	dairy processing	meat processing
Location	Western-Hungary	Western-Hungary	Western-Hungary	Eastern-Hungary
Interviewee	plant manager	farm manager	plant manager	CEO

Source: own construction

Table 2: Interview data.

by the fact that although the number of interviews is low, they were conducted with the three largest subsectors of the Hungarian food industry, the meat industry, the dairy industry, and the bakery and pasta industry. Three of the four firms are market leaders in their sub-sector, but because of confidentiality reasons, we cannot reveal them.

Based on these, our experience with the digitisation of the food industry in Hungary is diverse. The wide range of technologies and solutions used in the automotive industry (Nagy, 2018; Demeter et al., 2020) is far from that of the food sector. At the same time, it is very important that many processes have long been robotized and/or automated in the food industry, otherwise, the efficiency and food safety that customers and the authority expect of companies would not be achieved. However, the further development of these, I4.0, digitalisation, and the exploitation of big data are in the focus of corporate thinking.

Results and discussion

In this section, we will examine companies' I4.0 practices according to the analytical framework created in Chapter 2.1.

1. Strategy and organisation

In the strategy and organisation analytical dimension, we examined phenomena such as the existence of the I4.0 strategy or the effect of I4.0 on strategy. We examined whether the organisational structure has changed as a result of digitalisation. It was also interesting where the initiation of improvements began.

For two of the four companies examined, the move towards I4.0 was induced by the growth strategy. With the capacities available before the developments, they were no longer able to further increase production, so a decision had to be made as to whether they were satisfied with the current size of their business or wanted to grow. The latter was chosen by both of them and the new technologies purchased include digital solutions.

With the investments made, the pasta factory has multiplied its output capacity, since it aimed to increase its domestic market share and conquering the export markets. We cannot talk about specific I4.0 strategy, but when building the new factory and selecting the technology, they clearly wanted to take advantage of the opportunities offered by digitalisation: "*The factory is 21st century standard, so far we have focused on it. Now, the coming years will be more about how we can support the back-office activities, which is also important.*" In the case of the company, organisational changes have not yet taken place, but they already see the need, because the number of employees in the IT and maintenance departments has started to grow significantly. High-volume developments - such as choosing the technology for the new factory - was a top management decision, but operators were also involved in the design, fine-tuning and minor improvements.

In the case of the dairy production company, in addition to growth, survival and regulatory compliance also contributed to the technology innovation decision. The plant was obsolete by the end of the 2000s, and it was not possible to grow or increase efficiency with the existing number of workforce and technology: "*Company decided to develop this plant*". Regulatory standards also tightened, so they definitely had to improve. To this end, "*a large-scale project was launched and one of the best and most efficient sites nationwide was established*" - as our interviewee told us. As a result of the development, it was also possible to collect data that could be used to further increase animal welfare. In this case, we cannot talk about a developed I4.0 strategy, nor have any organisational changes been reported.

The dairy processing company is a member of an international group with a well-defined I4.0 strategy, and the Hungarian subsidiary also has responsibilities to fulfil it. According to the interviews, we distinguish between group-level developments that focus primarily on big data and analysis, improving forecasting and planning accuracy; and smaller, local-level developments that utilize I4.0 and digitalisation solutions to address a specific enterprise problem. "*We realize our ideas and a lot of things that have already been implemented in other factories in the group, we take over here.*" The most common force that triggers local development is the lack of manpower, which they try to solve with partial automation and robotization. The company, through the projects implemented, is explicitly seeking to increase food safety and to meet or exceed standards. The C-level director of digital development is present at the international level only, in the Hungarian factory there is no dedicated leader. Local development can come from any level of the organisation, and operators are motivated to submit their proposals.

For a meat processing company, there is no explicit I4.0 strategy as well: "I*t will affect* *the industry within 5 to 10 years (I4.0). Now it's only in the minds of those executives who are open.*" In their case, the developments began as a result of a risk analysis that revealed the jobs where the risk of a labour migration would seriously jeopardize the operation of the plant. Subsequently, steps were taken to automate and robotize these operations. The other direction of development when building a new plant was sensory data collection, monitoring of the meat processing process, and thus increasing process efficiency. The sources of development proposals are mainly senior management, but employees at the bottom of the organisational hierarchy also have the opportunity to submit ideas. Organisational changes have not yet occurred due to digitalisation.

We have found that the companies are changing, but in all cases, the digital transformation is underway, the starting point is mainly top management decided to move the company towards growth and efficiency. The specialty of the food industry is that better compliance with regulations also motivates the developments, as it was suggested by Kittipanya-ngam and Tan (2020) and Carpenter and Wyman (2017). We also found that, although improvements had begun, this did not change the strategy, there was no formal I4.0 strategy and the transformation of the organisation to the new business model had not yet begun. Amongst the reasons for the development of changing customer needs and retailers, the pressure was not identified in the case companies, contrary to literature.

2. Data collection and analysis

The purpose of this section is to explore what data the companies collect, and how it is collected, processed, and used in decision making.

The pasta company has intended to take advantage of the technology when designing the new production facility and gathers as much data as possible about the production process. The raw material entering the automatic production line passes through sensors of weight and humidity. The consistency of the dough being made is also controlled and so is the temperature of the dryer. The finished pasta passes through several control points both before and after packaging, where any piece of metal found in the production process is screened with a metal detector and the weight and bar code of the package are also checked. The machines are equipped with preventive maintenance sensors (operating hours meter), any deviation from the norm can be detected immediately and the necessary intervention can be decided. The manufacturer also provides remote diagnostic services and uploads maintenance tasks to be performed when a particular operating hour is reached. The goal is to reduce downtime or, if there is a downtime, to find the causes and avoid them next time. Based on these, the maintenance task can be scheduled, employees receive the task on a tablet, which informs them what task to perform, what materials to use, what error to correct (or not if it failed). There is also a monthly report on the machine's performance and operating characteristics. Analysing data is a big task for a company. By their own admission, they did not expect the analysis of the extracted data to be so important when building the new plant, so the development of the supportive IT background is lagging. Their goal is to have the most important data instantly online and make real-time decisions: "*So if the production manager wants to look at the production, then he doesn't have to look at the machine and see what happens, he just takes out his notebook and can watch the process.*"

The technology of the *dairy production company* enables the collection of a large amount of data. At starting the dairy production process, animals are individually identified using an RFID chip hanging in their ears. This chip stores the animal's identity, breed, ancestry, physical characteristics (e.g. lactation phase), and all interventions in which the animal was involved (vaccinations, calving, etc.). Milk yields from individual cows are also known, so when the animal arrives at the milking machine, it calculates an expected milk yield: "*If the cows at the start of lactation, then obviously the expected amount of milk is up, if she is the end of lactation the milk production will keep going down then the system obviously will change the expectations day by day. We get a lot of data from this device.*" If the expected milk yield is not met, it will be checked first if the milking machine has been set up correctly, and if the yield is less than expected for an extended time period, the animal will be examined by a veterinarian. The milking machine also checks the quality and composition of the milk that is being milked. It can calculate from the milk flow rate whether the milking worker has properly triggered the milking reflex (that is, the animal has given as much milk as possible), which is also reflected in the performance pay. The milking machine allows access to a variety of reports. The milking process can be queried in different compositions and focus. Latter are determined and evaluated by the animal breeding engineer. The availability of preventive maintenance and remote diagnostics was an important consideration when choosing the technology. The manufacturer constantly monitors the condition of the machines and makes recommendations for necessary maintenance. The feed consumed by the animals is also recorded in the computerized system. Cows are grouped according to their age, milk yield, lactation status, and receive a feed of appropriate quality and quantity. The composition of the fodder needed for a group of animals is determined by the feeding manager who uploads it into a flash drive and from that onto the feed mixing machine, which indicates which type of feed to load (by weight) and then mixes them.

Digital development, encouraged by the parent company of the *dairy processing firm*, is data collection and analysis. They are already trying to extract and analyse data from the PLC (Programmable Logic Controller). This can be a great help in preventive maintenance or even avoiding quality problems e.g., the scales built into the production process can monitor which component is being dosed and, if out of range, suggest which component should be added more, which is less. Tracing allergenic substances is also a requirement for food safety: "*In our industry, traceability of batches is very important, especially if the product contains allergenic substances.*" The aim is to see production data that allows immediate intervention.

Developments that ensure food safety and compliance with official regulations are of paramount importance to the company, as the internal rules of the international group are stricter than Hungarian or even EU rules. For example, when a component is added to a product, it is recorded in which raw material is blended from which supplier's production batch, thus ensuring traceability. This tracking is even more thorough with allergenic materials when it is also recorded who cleaned the machine after production.

The *meat processing company* deliberately chose digital technology to monitor processes and closely control the production line at its new plant, thereby helping to increase the accuracy of data control and process efficiency, and to support performance pay: "*we built in everything that can provide data for ERP, controlling and process control*". Another area of data collection is the monitoring of machine conditions for preventive maintenance. Machine shutdown is a major cost factor in this sector as well, and an important reason for the development of data collection in this area.

Collecting and analysing data was a very important factor in every company when choosing technology. Companies also report that they have a wealth of data and are struggling to process it and turn it into relevant information that they can use in real-time decision-making. At the same time, it seems that the production and storage of data is just one problem, the other is the development and purchase of software and systems that can analyse and display them. Food safety - as a reason behind developments - came up several times during interviews, as the literature claims (Kittipanya-ngam and Tan, 2020). However, the literature does not address the cost and challenge of building and operating such IT systems at companies that can store and share data of quality and composition compliant with food safety regulations. Although blockchain would be a suitable technology for solving this problem, we have not identified any traces of the application in the Hungarian company practice, yet. An appropriate system for data collection and analysis is challenging in all respects and companies are investing significant sums in this regard.

3. Technological solutions

In this section, we present all the technological solutions - be it automation, robotics, or IT development - that have resulted in great progress in the company in terms of both efficiency and data processing. We also cover the issues of data security and cyber security that concern companies.

The *pasta company* is implementing a high degree of automation and robotization in its new factory. The production of the dough goes through an automatic process from the breaking of the eggs to the unloading of the finished and packed pasta in the form of a unit load. In the dough-making machine, sensors monitor the dough's texture, moisture, humidity, and temperature. Heat recovery is built in the drying process, which allows significant energy savings. The packaging is carried out by an automatic packaging line after production. After several weighing and metal detector screening, the pasta packages are placed in a cardboard box and then conveyed by the conveyor to the warehouse preparation area. Here, unit-loading robots receive the boxes and form stacks of products of the same type identified by barcode scanning. The stacks are captured by the wrapping machine and then, through it, an automatic signal is sent to the AGV (automatic guided vehicle) forklifts, which, after identification, transfer the unit load to the high bay racks warehouse designated for end products and hand it over to the high bay racking machine. When the unloading request is received, the former process is reversed, the racking machine finds the desired unit load, hands it over to the automatic forklift, and it forwards it to the delivery area.

Sensors built into the above process (balance, barcode scanner, metal detector, etc.) continuously produce data that is stored on company servers. The plant manager highlighted cyber security, which they try to secure through operating their own server and differentiating permissions.

When purchasing the technology, the possibility of connecting the machines and equipment to each other and the production centre was especially important. Remote diagnostics allow the company to detect the fault much faster, spare parts can be ordered sooner and, in the case of foreign-made machines, does not incur the cost of transporting the service staff to the site: "*Each machine has a remote diagnostic service and the device monitors continuously the operation, forces, temperatures, etc. In case the system signals or detects a malfunction or abrasion, we will be notified and then obviously we can request proper service for the machine*".

In the case of a *dairy producing company*, it is worth examining two directions of technology, on the one hand, milking technology, and on the other hand, animal welfare equipment. They identified cattle by RFID chips hanging in their ears, which store plenty of data about them (see chapter Results and Discussion, Data analysis). The milking machine is set up manually, but the milking machine milks the cattle in 4-5 minutes and automatically stores quantitative and qualitative information: "*The computer automatically calculates the expected milk from the previous days' milk production and lactation stage*". Animal welfare fundamentally affects milk yield, so several "convenience" solutions can be observed at the farm. For cattle, temperatures above 25 degrees Celsius cause heat stress, so above this temperature limit, the water spray is switched on in the paddock, which automatically releases water more and more frequently as the temperature rises. Similarly, when the temperature limit is reached, the fans also turn on. The basically open side-walled paddocks allow good airflow, but in case of bad weather (rain, wind), a tarpaulin automatically lowers to protect the animals from discomfort.

The computer-determined feed composition is mixed in the feed mixer, which indicates how much nutrients should be added to each group of goods according to the uploaded plan. The fodder is poured along with the paddocks and can be consumed by the cows leaning out of the paddock. While eating, they inevitably break through the paddock, so a feed-sweeping robot walks through the paddocks several times during the day to dig it back to where the animals can reach it. The content of the feed is checked several times a week and the composition is revised if necessary.

At the *dairy-producing company*, the technology is not as integrated as in the pasta factory. The investments made focused on key processes to comply with food safety regulations which would not be possible otherwise. The system stores a lot of information about animals, skimmed milk, a significant part of which is prescriptive, while analysing the other part of the data allows creating an appropriate environment, nutrients, and caring for the animals. The data collected is used to increase animal welfare, formulate optimal feed, and determine workers 'performance pay. Although there is a significant amount of data generated that is stored on corporate servers, the issue of cybersecurity has not been addressed.

The *dairy processing company* makes developments in two directions. On the one hand, there are expectations from the parent company to improve productivity and carry out other digitalisation projects, and on the other hand, employees within the company are very creative and make suggestions for improvement. This company operates a very labour-intensive production process, which resource, however, is scarce. To replace the workforce, cobots (cooperative robots) have been implemented that can be operated safely among humans and take on monotonous, demanding or repetitive tasks: "*Cobots are relatively smaller and therefore cheaper than large industrial robots and can be implemented amongst humans. Work safety is very important*". A similar palletizing robot works as at the pasta manufacturer, which is not an own, customized development, a ready-made system has been purchased and adapted. Like the dairy producer, a lot of data is collected (Results and discussion, Data analysis)) and processed, some of which are used to fit regulations and others for process development and efficiency. The data is stored partly locally and partly at a group level at central data storage, while analyses are carried out at all levels. The goal is to be able to intervene immediately as soon as an emergency warning is generated.

The new plant at the *meat processing company* has been specifically built to give management a complete overview of the process through built-in sensors. Here, the process of raw material arrival, receiving, and first processing is monitored, which are essential for the resale of the raw material or processing by further plants. Based on the sensor data, the net expected amount of gross incoming raw material can be better planned, the system not only calculates the loss (gross-net deviation) but also monitors the raw material weight, moisture, and protein content: "*We use sensors in the raw material processing. This is repeated at several stages during the process*." Once again, the goal is the possibility of immediate intervention. Although they cannot yet measure efficiency per worker, only at the production line level, they can already use the data for performance pay. Another element of development was the equipment of machines with preventive maintenance functions to reduce downtime. The company used lean solutions before, and the current developments were also coordinated by this team.

Overall, in connection with the development of technology, most companies find proper digital solutions, the connection of machines to the network is realized, data collection works, it is partly analysed and used in some part of decision-making. An I4.0 solution can be observed in the pasta factory, in which machines are not only connected to each other, but the production forms a unified whole with the automated warehouse and logistics processes, and automated, robotic processes can also be observed. Besides digitalisation, robotics and automation are already appearing at the milk processing company, but the solutions do not form a system, isolated developments can be observed.

There are many European Union and Hungarian state funds available for technological developments. The majority of companies examined have won tenders and have also been able to achieve a favourable return on their deductible. However, participation in further tenders is now so expensive because of the price of the technology and the construction, and so difficult to find a contractor for the project, that these jeopardize the project to fit in the budget and the time frame at all.

4. Employees

The purpose of analysing this dimension is to examine how the company involves its employees in digital developments. We also examined what jobs were created and transformed as a result of new technologies and process developments.

Contrary to what is described in the literature (Deloitte, 2015; Carpenter and Wyman, 2017), the examined companies not only do not decrease staff due to technological developments, but rather struggle with labour shortages, or, only improvements can help to avoid shortage.

The *pasta company* has always used a lot of automation in production, so the blue-collar staff was low in the past, too. Currently, the pasta factory employs approximately 100 people, producing more than 50,000 tons of pasta per year, in three shifts a day. As a result of the current developments, they realize that they will increasingly need workers who can operate, maintain, or program machines. In this respect, they have no problems yet, but the increase in staff has begun. The number of maintenance staff has doubled in the last 5 years and the organisation of the technical directorate has also expanded. In total, 14 people are already working to ensure the technical maintenance of production. Because automation was already present in production processes in the past, accepting the new technology with workers was not a problem, although generational differences appear (young people are basically easier to get used to working with a robot, they also learn programming faster than their older peers). "*By automating a lot of things, we've triggered a lot of manual work. On the other hand, there is a great need for a staff that can not only operate these machines but also understands them and intervene if necessary*" the interviewee said. Workers on the production line are also involved in generating development ideas, and company management is open to development proposals from any level.

The *dairy production company* is faced with the situation that the improvements made it possible that they have enough manpower: "*There is a growing shortage of manpower here, and unfortunately I have to say that it is not a question of money. Because if I were to pay twice as much, there would certainly not be more milkers who want to do it.*" The roles of veterinarians, animal husbandry engineer, and feed manager have been expanded through digitisation, they keep electronic records, upload animal data to the central register, compile reports that summarise and analyse milking results, and electronically plan the fodder, which is assembled and loaded into a feed mixer. Their task is to query and compile new types of reports from the available data, which can be further developed. At the company, development ideas come more from white-collar employees, and investment planning is a top management decision. The company also faced the dilemma of knowing that more advanced technology would be available, and they would buy it, but there are very few professionals in the country who could be able to operate it. This is seen as a serious constraint of further development.

The *dairy processing company* is struggling with labour shortages. The unemployment rate in the region is very low (1-2%), so they often have to hire workers they consider just right to do the job, and the fluctuation is high. Its developments have also been driven partly by the creation or assistance of jobs in areas with the greatest labour shortages and fluctuation. The workers accept the technology, are happy to work with the cobots, and experience it as a modern technological environment. On the one hand, the company receives the development directions from the parent company, on the other hand, generates them from the inside. The source of the latter is typically engineering, quality assurance, or plant management. However, in the case of a problem, they first look at the group level, whether a similar one has arisen somewhere amongst the affiliates, and how it has been solved there: "*Colleagues here are very creative, there are also Western European factories asking for advice from us*". For the time being, the company does not feel that the structure of the workforce is changing due to digitalisation and robotization, but they see that in 5 years they will need many more technicians, PLC programmers, and maintainers.

For the *meat processing company*, the labour shortage is also a typical problem. There are jobs in which learning manual work is very time-consuming, and with the knowledge gained, many are more likely to be move to Western Europe. That is why one of the drivers of the developments is the replacement of jobs where there is a high risk of emigration. However, there are complex manual tasks that cannot be taught to a robot: "*I see only partial opportunities (for the introduction of I4.0 devices) in production and more in logistics especially in warehousing. It is also important in operations and maintenance.*"

Sources of development ideas can be both blue-collar and white-collar workers, and the company is open to initiatives, and proposals are rewarded if they are successful.

Based on the above, it can be seen that companies perceive that digitalisation is changing expectations of the workforce, making jobs more complex, and requiring much more technological knowledge. However, labour shortage affects food companies seriously, the suction power of other sectors is very strong.

5. Company culture

In examining corporate culture, we looked at how open companies are to innovation, how well their employees have the opportunity to make suggestions, and how knowledge sharing takes place in the company.

Although the implementation of I4.0-related development was partly forced for most of the companies, the firms made very important and significant investments. To maintain growth, the *pasta company* has embarked on the development of a new, modern production unit, which has made it a significant player not only in Hungary but also in Central-Eastern Europe. The commitment of senior management and the involvement of employees should be highlighted as a factor of success. When choosing the technology, it was a clear consideration to choose one, which can serve competitiveness, not to seek the cheapest solution. An explicit knowledge management system does not work t the company, but since it is a small organisation, the communication takes place directly: "*Training has always been an important aspect here in the pasta factory.*"

The *dairy producing company* opted for a large-scale and fundamental technological change to survive. The decision was made by the top management, but like the pasta company, the management mobilized great energies so that the decision-makers of the investment could learn about the technological possibilities. They went on study trips abroad to get to know and learn about the technology, so they sacrificed a lot to make an informed decision and to prepare for the operation: "*We constantly - with the support of the top management - go on study trips and employ consultants who are currently introducing new things, novelties*". In terms of corporate culture, employees are expected to be constantly learning and open to change.

The *dairy processing company* has taken many solutions from the parent company. Innovation is very important centrally, and central digital developments must be carried out. At the group level, a knowledge-sharing portal is now being set up, in which subsidiaries upload the solution to a specific problem, their best practices. They make common knowledge for the group, which can be used by any other subsidiary. Bottom-up initiatives have a culture in local development: "*Engineering brings together (the ideas), designs a solution for it, calculates how much it would cost, what will improve, how long it can be realized, what savings can be expected*".

The *meat company* has also taken great strides towards modernization, economic stability, and technological development. According to the expert interviewed, it is very important to have a management in a company that has a vision and can assess where the industry is heading over the next 10 years and dares to make the right decisions and takes risks. As the interviewee says: "*Here you have to decide not only based on return, but there will be other aspects as well, e.g., where the industry will be in 10 years, what consumer needs will be and what technology will be able to meet it*". In recent years, a culture of acceptance of bottom-up initiatives has also developed in this organisation, but we cannot talk about a knowledge-sharing system.

In the previous chapters, we got to know four organisations committed to innovation and embarking on change. Corporate culture also supports development initiatives in all places, whether they come from below or from above, but the operation of a knowledge management system - except for a company with an international background - is not typical.

Conclusion

In our study, we examined the steps taken by the food industry towards digitalisation and I4.0. We sought to increase the validity of the research by interviewing actors from the three most important sub-sectors of the Hungarian food economy.

The main finding of the study is that there are food industry developments in Hungary that take advantage of the opportunities offered by digitalisation and I4.0. Table 3 is following the same structure and logic we created within our analytical framework. Related to strategy and organisation dimension, two main questions were asked. On the one hand, whether the I4.0 strategy exists or not, and on the other hand if the need for organisational changes is obvious for the management and already started. As the table shows, only the dairy processing company has an I4.0 strategy (on an international level due to its membership of an international group), and only the pasta manufacturer recognized the need for organisational changes due to the new positions (IT, maintenance). Data collection and data analysis questions are raised in the second dimension, where we could observe the developing status of the companies. Within the technology dimension, we distinguished the field of solutions applied. It came as a question, whether the whole process integrates I4.0 technologies or only it can be found partially, in isolated solutions or supportive processes. In connection with employees' dimension, we concluded that usage of I4.0 technologies in none of the cases ended up with less need for human work but on the contrary. It helped to decrease the risk due to the human workforce shortage. Also, the need for new positions (maintenance, IT) appeared in our cases. Related to the involvement of the employees we observed an in-progress status of the companies, as the developments, ideas come mainly from the top management. Our last dimension is the company culture with two areas: the openness for innovation and knowledge sharing. We found that all our cases belong to an in-progress openness for innovation as the food industry is typically follower compared to very innovative industries. The knowledge-sharing aspect of the companies differed significantly in our research. In two cases, formal or informal levels of knowledge sharing practices can be observed on group levels. However, in the other two cases, knowledge sharing is reaching only a basic level as it is informal and occasional. The main results are summarised in Table 3.

According to Kürthy et al. (2016), the Hungarian food industry is lagging behind its Western competitors in terms of efficiency and technology, although export activity significantly boosts the sales revenue of enterprises. This is also evident in the pasta factory which has reached its capacity limit and has committed itself to a growth strategy that it plans to build on exports by setting up a new plant. To achieve higher efficiency, companies can best help to implement further optimisation and increase their competitiveness by purchasing state-of-the-art technology and analysing the data it produces. Consistent

Results		*Pasta manufacturer*	*Dairy producer*	*Dairy processor*	*Meat processor*
Strategy and organization	I4.0 strategy	No	No	Yes, on international level	No
	Organizational changes	Embryonic	No	No	No
Data analysis	Data collection	Extensive	Extensive	Developing	Limited
	Data analysis	Developing	Developing	Developing	Developing
Technology	Automation, robotization	Integrated process	Isolated solutions	Isolated solutions	In supportive processes
Employees	Cadre personnel	Increasing in maintenance	Shortage	Shortage	Shortage
	Involvement	In progress	No	In progress	In progress
Company culture	Openness for innovation	In progress	In progress	In progress	In progress
	Knowledge sharing	Group level, informal	Basic	Formalized, on international level	Basic

Source: own editing

Table 3: Main findings of the study.

with the Simutech (2016) survey, interviewees reported significant investments (pasta, milk producer, and meat processor) in which decision was based not on the price of the technology rather on its knowledge and capabilities. One of the most important functions is predictive maintenance, data processing, in which they are still working on the exploitation of opportunities. The low prevalence of I4.0 is indicated by the fact that organisational changes were not typical at the companies, but rather the increase in the number of technical and maintenance departments.

Continuous quality control and improvement is also a central element of corporate development (Santeramo et al., 2018). Productivity has improved, enabling customers to access a wider range of products at affordable prices (pasta), improving and controlling product content (milk production, meat processing), and paying attention to allergens at a higher level (dairy processor).

The results of our study are completely consistent with the research of Bibi et al. (2017), Carpenter and Wyman (2017), and Bottani and Rizzi (2008), according to which the development of technology is an important tool for complying with always tightening regulations. Barcodes play a primary role in tracking, but RFID chips are used to identify and track cows at the dairy farm RFID (or barcode) -based traceability covering the entire supply chain as Tian (2016) suggest is not yet available, and supply chain actors are still looking for a solution to the problem individually. In the same way, we could not find evidence of applying (or at least planning to apply) blockchain technology in the analysed firms.

Our research revealed the effects of digitalisation on the food industry companies. We highlighted that besides the technological development firms have to adapt to their organisation, as well, and the role of human factors in success is essential. Success does not depend only on the technology acceptance of blue-collar workers, but also on their development ideas as well as the top-management engagement. Our important finding is that although food security regulations force companies to technology – especially IT – development, and besides the high cost it is also challenging in terms of data security, storage, and processing capacity and methodology.

In our exploratory study, four corporate practices were introduced which also can serve as excellent benchmarks. We highlighted that although I4.0 and digitisation appear in various ways in the different companies and sub-sectors, their impact is far from limited to production, affecting several levels of organisational operation.

The research also has limitations. The case study method and the four analysed companies do not allow generalising the results, although it is excellent in revealing how digitalisation affects companies and aims to gain a better understanding of *what* and *why* firms decide and do to move towards I4.0. Regarding the future research prospects, we are intended to track the companies further and follow up their ways on the road of digitalisation. We are eager to know how they continue the development and how their organisation will transform thank to I4.0.

Acknowledgments

This research was supported by the project "Aspects on the development of intelligent, sustainable, and inclusive society: social, technological, innovation networks in employment and digital economy" (EFOP-3.6.2-16-2017-00007). The project has been supported by the EU, co-financed by the European Social Fund and the budget of Hungary.

Corresponding authors
Judit Nagy, PhD, associate professor
Corvinus University of Budapest, 8 Fővám tér, 1093 Budapest, Hungary
Phone: +36 1 482 5232, E-mail: judit.nagy@uni-corvinus.hu

References

[1] Bibi, F., Guillaume, C., Gontard, N. and Sorli, B. (2017) "A review: RFID technology having sensing aptitudes for food industry and their contribution to tracking and monitoring of food products", *Trends in Food Science & Technology*, Vol. 62, pp. 91-103. ISSN 0924-2244. DOI 10.1016/j.tifs.2017.01.013.

[2] Bottani, E. and Rizzi, A. (2008) "Economical assessment of the impact of RFID technology and EPC system on the fast-moving consumer goods supply chain", *International Journal of Production Economics,* Vol. 112, No. 2, pp. 548-569. ISSN 0925-5273. DOI 10.1016/j.ijpe.2007.05.007.

[3] Brettel, M., Friederichsen, N., Keller, M. and Rosenberg, M. (2014) "How virtualization, decentralization and network building change the manufacturing landscape: An Industry 4.0 Perspective", *International Journal of Mechanical, Industrial Science and Engineering*, Vol. 8, No. 1, pp. 37-44. ISSN 0020-7403.

[4] Carlozo, L. (2017) "What is blockchain?", *Journal of Accountancy,* Vol. 224, No. 1, p. 29. ISSN 1945-0729.

[5] Carpenter, G. and Wyman, O. (2017) "F*ood manufacturing – Are you ready for Industry 4.0?*", Marsh Report, Marsh and McLennan Companies, United Kingdom [Online]. Available: https://www.marsh.com/uk/insights/research/food-manufacturing-are-you-ready-for-industry.html [Accessed: 5 June 2019].

[6] Chaime, J. (2020) "*World population: 2020 overview*". [Online]. Available: https://yaleglobal.yale.edu/content/world-population-2020-overview [Accessed: 29 April 2020].

[7] Dalenogare, L. S., Benitez, G. B., Ayala, N. F. and Frank, A. G. (2018) "The expected contribution of Industry 4.0 technologies for industrial performance", *International Journal of Production Economics*, Vol. 204, pp. 383-394. ISSN 0925-5273. DOI 10.1016/j.ijpe.2018.08.019.

[8] Deloitte (2015) "*Industry 4.0 – An introduction*", Deloitte, Hollandia [Online]. Available: https://www2.deloitte.com/content/dam/Deloitte/be/Documents/Operations/2015_Industry%204%200%20Report%20vFinal.pdf [Accessed: 20 April 2019].

[9] Demartini, M., Pinna, C., Tonelli, F., Terzi, S., Sansone, C. and Testa, C. (2018) "Food industry digitalization: from challenges and trends to opportunities and solutions", *IFAC-PapersOnLine*, Vol. 51, No. 11, pp. 1371-1378. ISSN 2405-8963. DOI 10.1016/j.ifacol.2018.08.337.

[10] Demeter, K., Losonci, D. and Nagy, J. (2020) "Road to digital manufacturing – A longitudinal case-based analysis", *Journal of Manufacturing Technology Management*, ISSN 1741-038X. DOI 10.1108/JMTM-06-2019-0226.

[11] Denzin, N. K. and Lincoln, Y. S. (Eds) (2011) "*The SAGE Handbook of Qualitative Research*", Sage, Thousand Oaks, USA. ISBN 978-1-4129-7417-2.

[12] Dworschak, B. and Zaiser, H. (2014) “Competences for cyber-physical systems in manufacturing– first findings and scenarios”, *Procedia CIRP*, Vol. 25, pp. 345-350. ISSN 2212-8271. DOI 10.1016/j.procir.2014.10.048.

[13] Fettig, K., Gačić, T., Köskal, A., Kühn, A. and Stuber, F. (2018) “Impact of Industry 4.0 on Organisational Structures”, In 2018 *IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC),* pp. 1-8. DOI 10.1109/ICE.2018.8436284.

[14] Frey, C. B. and Osborne, M. A. (2017) “The future of employment: how susceptible are jobs to computerisation?”, *Technological Forecasting and Social Change,* Vol. 114, pp. 254-280. ISSN 0040-1625. DOI 10.1016/j.techfore.2016.08.019.

[15] Gauger, C., Gehres, B., Quinn, M., Schmieg, F. and Xu, G. (2017) “Building the digital car company of the future”, The Boston Consulting Group, USA, April.

[16] Geissbauer, R., Vedso, J. and Schrauf, S. (2016) “*Industry 4.0: Building the digital enterprise*”. [Online]. Available: https://www.pwc.com/gx/en/industries/industries-4.0/landing-page/industry-4.0-building-your-digital-enterprise-april-2016.pdf [Accessed: 20 April 2019].

[17] Gill, M. and VanBoskirk, S. (2016) “The digital maturity model 4.0. Benchmarks: Digital Transformation Playbook”.

[18] Gracel, J. and Łebkowski, P. (2018) “The Concept of Industry 4.0-Related Manufacturing Technology Maturity Model (ManuTech Maturity Model–MTMM)”, *Decision Making in Manufacturing and Services*, Vol. 12, No. 1-2, pp. 17-31. ISSN 2300-7087. DOI 10.7494/dmms.2018.12.1-2.17.

[19] Hermann, M., Pentek, T., and Otto, B. (2016) “Design principles for industrie 4.0 scenarios”, In 2016 *49th Hawaii international conference on system sciences (HICSS)*, pp. 3928-3937. E-ISBN 978-0-7695-5670-3. ISSN 1530-1605. DOI 10.1109/HICSS.2016.488.

[20] Hofmann, E. and Rüsch, M. (2017) “Industry 4.0 and the current status as well as future prospects on logistics”, *Computers in Industry*, Vol. 89, pp. 23-34. ISSN 0166-3615. DOI 10.1016/j.compind.2017.04.002.

[21] Horváth, D. and Szabó, R. Z. (2019) “Driving forces and barriers of Industry 4.0: Do multinational and small and medium-sized companies have equal opportunities?”, *Technological Forecasting and Social Change*, Vol. 146, pp. 119-132. ISSN 0040-1625. DOI 10.1016/j.techfore.2019.05.021.

[22] Ilie-Zudor, E., Kemény, Z., Van Blommestein, F., Monostori, L. and Van Der Meulen, A. (2011) “A survey of applications and requirements of unique identification systems and RFID techniques”, *Computers in Industry,* Vol. 62, No. 3, pp. 227-252. ISSN 0166-3615. DOI 10.1016/j.compind.2010.10.004.

[23] Kittipanya-ngam, P. and Tan, K. H. (2020) “A framework for food supply chain digitalization: lessons from Thailand”, *Production Planning & Control*, Vol. 31, No. 2-3, pp. 158-172. E-ISSN 1366-5871. ISSN 0953-7287. DOI 10.1080/09537287.2019.1631462.

[24] Kürthy, G., Dudás, G. and Felkai, B. O. (Eds.) (2016) “A magyarországi élelmiszeripar helyzete és jövőképe”(The current situation and the future of the Hungarian food industry), Agrárgazdasági Kutató Intézet, Budapest. ISBN 978-963-491-599-7. (in Hungarian).

[25] Lichtblau, K., Stich, V., Bertenrath, R., Blum, M., Bleider, M., Millack, A. and Schröter, M. (2015) “IMPULS-industrie 4.0-readiness”, Impuls-Stiftung des VDMA, Aachen-Köln.

[26] Nagy, J., Oláh, J., Erdei, E., Máté, D. and Popp, J. (2018) “The role and impact of industry 4.0 and the internet of things on the business strategy of the value chain— The case of Hungary”, *Sustainability*, Vol. 10, No. 10, p. 3491. ISSN 2071-1050. DOI 10.3390/su10103491.

[27] Pagani, M. and Pardo, C. (2017) “The impact of digital technology on relationships in a business network”, *Industrial Marketing Management*, Vol. 67, pp. 185-192. ISSN 0019-8501. DOI 10.1016/j.indmarman.2017.08.009.

[28] Santeramo, F. G., Carlucci, D., De Devitiis, B., Seccia, A., Stasi, A., Viscecchia, R. and Nardone, G. (2018) "Emerging trends in European food, diets and food industry", *Food Research International*, Vol. 104, pp. 39-47. ISSN 0963-9969. DOI 10.1016/j.foodres.2017.10.039.

[29] Schuh, G., Anderl, R., Gausemeier, J., Ten Hompel, M. and Wahlster, W. (Eds.). (2017) "Industrie 4.0 Maturity Index - Managing the Digital Transformation of Companies", Munich: Herbert Utz.

[30] Schumacher, A., Erol, S. and Sihn, W. (2016) "A maturity model for assessing Industry 4.0 readiness and maturity of manufacturing enterprises", *Procedia Cirp*, Vol. 52, pp. 161-166. ISSN 2212-8271. DOI 10.1016/j.procir.2016.07.040.

[31] Sen, K., Durakbasa, N., Erdol, H., Berber, T., Bas, G. and Sevik, U. (2017) "Implementation of digitalization in food industry", *DAAAM International Scientific Book,* pp. 91-104. ISSN 1726-9687. DOI 10.2507/daaam.scibook.2017.08.

[32] Simutech (2016) "*Industry 4.0: How the food industry must adapt to survive*", Simutech Multimedia, White Paper, Kanada. [Online]. Available: https://www.simutechmultimedia.com/downloads/whitepapers/simutech-whitepaper-industry-4.0.pdf [Accessed: 5. June 2019].

[33] Tian, F. (2016) "An agri-food supply chain traceability system for China based on RFID & blockchain technology", In 2016 *13th international conference on service systems and service management (ICSSSM)*, pp. 1-6. E-ISSN 2161-1904. DOI 10.1109/ICSSSM.2016.7538424.

[34] Tse, D., Zhang, B., Yang, Y., Cheng, C. and Mu, H. (2017) "Blockchain application in food supply information security", In 2017 *IEEE International Conference on Industrial Engineering and Engineering Management (IEEM),* pp. 1357-1361. E-ISSN 2157-362X. DOI 10.1109/IEEM.2017.8290114.

[35] World Economic Forum (2018) "*The system initiative on shaping the future of food security and agriculture*". [Online]. Available: https://weforum.ent.box.com/v/FSA [Accessed: 26 Feb. 2019].

[36] Yin, R. K. (2011) "*Applications of case study research*", Sage Publications, USA, pp. 4-6. ISBN 978-1-4129-8916-9.

Animal Husbandry Export Measures Productivity: What is the Position of the Czech Republic?

Lenka Rumánková[1], Elena Kuzmenko[1], Irena Benešová[1], Luboš Smutka[1], Adriana Laputková[2]

[1] Department of Economics, Faculty of Economics and Management, Czech University of Life Sciences Prague, Czech Republic

[2] Institute of British and American Studies, Faculty of Arts, University of Presov, Slovakia

Abstract

Efficiency, productivity and competitiveness are important indicators of any production process. Speaking of agriculture and, in particular, animal husbandry in the light of sustainable development context, the values of these indicators become of special importance. Since adequate and efficient usage of disposable labour, land and capital can play a crucial role in obtaining and preserving higher efficiency and productivity levels, the natural question arises - does any relationship exist between mentioned production factors and revealed comparative advantage?

The main aim of this research is to evaluate the relationship between export-measured productivity and comparative advantages in animal husbandry of selected European countries. The benchmark is provided in relation to the Czech Republic. To analyse productivity of agricultural and, more specifically, animal production in the European countries selected for the analysis (based on available balanced data incorporating the period from 2005 to 2017), a decision was made to trace export performance of these countries recalculated with regard to a unit of core productive factors, such as land, labour and capital. Based on the foreign trade indicators (Gruber Lloyd index, RCA index), cluster analysis was conducted, in which individual calculation was used as an input variable. Subsequently, hierarchical clustering and Ward´s method were used. The evidence from this study suggests that the revealed comparative advantage of the countries is not determined primarily by the level of export-based productivity. The relationship between these variables is rather weak and very often negative, which indicates that productivity indicators do not play a significant role in the overall competitiveness of the monitored countries.

Keywords

Productivity, competitiveness, comparative advantage, animal husbandry, export, European Union, Czech Republic, benchmark, selected countries.

Rumánková, L., Kuzmenko, E., Benešová, I., Smutka, L. and Laputková, A. (2020) “Animal Husbandry Export Measures Productivity: What is the Position of the Czech Republic?", *AGRIS on-line Papers in Economics and Informatics*, Vol. 12, No. 4, pp. 93-110. ISSN 1804-1930. DOI 10.7160/aol.2020.120407.

Introduction

Efficiency, productivity, and competitiveness are important indicators related to evaluation of labour, land, capital, production, export, and other elements. The question is whether any (concrete) relationship among these indicators can be detected. It might be assumed that productivity can be considered as an important determinant and prerequisite for competitiveness. For this reason, this paper discusses the question in the field of agri-food export of selected EU countries.

According to, for example, Latruffe (2010), competitiveness should be measured with respect to a benchmark as it is a relative concept. Firms must be compared with each other, or nations with each other. Producing absolute figures for a country or an industry seems meaningless. Thus, the relationship between efficiency in terms of productivity and competitiveness of agri-food export in selected EU countries is examined and presented in this paper. Moreover, currently many approaches to examination of efficiency, productivity and competitiveness at different levels can be found. In order to meet the main objective of this paper, a suitable method had to be identified and employed.

Fojtíková and Staníčková (2017) analyse

export competitiveness and productivity of EU member states using the Factor Analysis and Data Envelopment Analysis (DEA). The results show that the level of export competitiveness is different in individual EU member states and that it changes during the analysed period (2000 – 2015). Similarly, Serrao (2003) employs the Data Envelopment Analysis (DEA) and the Stochastic Frontier Analysis (SFA) to examine the sources of agricultural productivity growth and productivity differences among individual countries and regions of EU in the analysed period (1980 – 1998). Pokrivčák et al. (2015) employ non-parametric DEA to examine the efficiency of the agricultural sector in EU-26 countries. Significant differences among EU countries have been detected in the analysed period (2007 – 2012). In conclusion, the efficient countries have relatively large farms and considerable expenditures for agriculture. Moreover, it has been concluded that the countries' productivity is strongly influenced by utilization of inputs. Rungsutiyawiboon and Lissitsa (2007) employ a parametric distance function approach to measure the Malmquist total factor productivity index. The results show that the transition countries (defined in the analysed period (1992 – 2002)) achieve higher agricultural performance. Serrao (2003) also employs DEA and Malmquist productivity index to examine the levels and trends in the global agricultural productivity of selected European countries. The sources of productivity growth over the time and the differences among the countries and the regions are defined in the analysed period (1980 – 1998).

Csikósová et al. (2018) examines the importance of work productivity in EU-28 as an important factor that influences economic growth and is also influenced by various determinants. Based on the research, rising differences among individual countries have been detected. Similarly, Rozkošová and Megyesiová (2017) define labour productivity as probably the main determinant of economic growth, the determinant affected by other factors. Based on the examination of EU-28, the differences among the individual states are considered in the analyzed period (2005 – 2016). Wulong and Beiling (2017) analyse effective multifactor productivity (MFP) growth in Canada, USA, Australia, Japan, and selected EU countries. The results show that the increase in effective MFP is closely related to the decline in output prices and improvement in international competitiveness.

Berthou et al. (2015) evaluate competitiveness, focusing on the relationship between the productivity and export performance among European economies. The research confirms that exporters are more productive than non-exporters and, additionally, the productivity premium is rising with the firms' export experience (permanent exporters are much more productive than starters). Moreover, what can be seen is that both the level and the growth of firm-level exports rise with the firm's productivity. Finally, it has been concluded that the shape of the productivity distribution within each country can have considerable implications in terms of the dynamics of aggregate trade patterns.

Bojnec and Fertö (2014) provide an insight into export competitiveness of meat products from EU-27 member states on the global market. The RCA index is used to analyse the level, composition, and evolution of the developmental patterns in the export competitiveness of meat products. The results show that except for some niche meat products, a large number of EU-27 member states experience a comparative disadvantage on global markets in the analysed period (2000 - 2011). Moreover, the revealed comparative advantage on the global markets are most robust for Ireland, Spain, the Netherlands, France, Belgium, Denmark, Poland, Cyprus, and Hungary. However, the RCA indices and their survival rates differ across the meat product groups. In conclusion, the heterogeneity in export competitiveness of EU-27 member states indicates the importance of the differentiation of meat products in competitive export specialization on global markets. Carraresi and Banterle (2008) examine EU competitiveness at the sector level on the intra-EU market. The analysis is conducted by assessing trade indices (RCA etc.). Moreover, cluster analysis is employed to classify groups of countries with similar features in terms of competitive performance in the analysed period (1991 – 2006). In conclusion, Spain is considered as the country attaining a high level of competitiveness. Contrariwise, the United Kingdom is detected as the country with the worst performance.

Galović et al. (2017) focus on international competitiveness of analysed countries through selected indicators. The results show that despite identical trade policy, external conditions and, for numerous EU member states, the same currency, trade performance of the member states is extremely diverse. Moreover, the most developed countries within the EU are consistent in their positive values and growth. These countries also have a string inclination towards the development, expansion, and competitiveness, and show no signs of stopping, given the positive trade balance. Bojnec and Fertö (2015) investigate competitiveness of agri-food exports of the EU-27 countries on global markets using the RCA index. The results show that in the analysed period (2000 - 2011) the majority of agri-food products in the EU-27 countries show a comparative disadvantage on global markets. It has been discovered that most old (EU-15) member states experience a greater number of agri-food products with longer duration of RCA than most new (EU-12) member states. The Netherlands, France and Spain are considered as the most successful member states in agri-food export competitiveness on global markets. Ružeková et al. (2020) assume that the higher quality of institutional environment is characterized by a higher level of competitiveness and lower transaction costs based on the belief that export performance is a reliable measure of competitiveness. However, the results demonstrate that export performance is not a universal indicator of competitiveness. Thus, it is necessary to apply other, especially multi-factor indicators. Furthermore, Nowak and Kaminska (2016) analyse competitiveness of EU-27 countries. Their research focuses on the relationship between production factors, productivity, and the importance of agriculture in international trade. In conclusion, based on the results for the analysed period (2009 -2011) the analysed countries are divided into four groups that are similar in terms of agricultural competitiveness.

The results of the examination of the productivity and export competitiveness in individual countries are presented, for example, in the following publications: Tiffin (2014) emphasizes the role of innovations on the export market share rather than price-based competitiveness. The high-quality export mix and the ability of small-scale specialized firms are considered as sources of strength of Italian export. Contrariwise, structural barriers that depress productivity have also been detected. Fertö and Hubbard (2003) examine competitiveness of Hungarian agriculture and food processing in relation to EU countries based on the indices of the revealed comparative advantage. The results indicate a comparative advantage in a range of agri-food products, including animals and meat, in the analysed period (1992 – 1998). Moreover, it has been concluded that the RCA indices are stable during the period of transition, although there is evidence of weakening in the comparative advantage level as revealed in the Balassa index. Gorton et al. (2000) analyse competitiveness of agricultural production in Bulgaria and the Czech Republic compared to international markets and EU. Their competitiveness is measured in terms of the Revealed Comparative Advantage (RCA) and Domestic Resource Cost (DRC) ratios. Among others, the results show that Czech livestock production is not competitive on the world market. Identical results are found in Bulgaria, with the exception of pork. Burianová (2010) employs the Balassa RCA index and Michaely index (MI) to analyse export performance of the Czech Republic. The results show that the commodities competitive on the EU market in the analysed period (2004 – 2008) can be found; moreover, measure of specialization is crucial in this evaluation.

Materials and methods

The main aim of this research is to evaluate the relationship between export-measured productivity and comparative advantages in animal husbandry of selected European countries. The benchmark is provided in relation to the Czech Republic (if not specified differently).

Based on the above facts, the research questions referring to the comparative advantages of the monitored countries and their position in the productivity can be formulated.

Research question 1: Comparative advantages of the monitored countries in individual commodity aggregates with regard to the productivity are currently similar.

Research question 2: The position of the Czech Republic with regard to export-based productivity is similar to that of other European countries.

Data description

To address the issue of export-measured productivity in selected EU countries representing the same climatic zone, raw data on exports in current USD prices were retrieved from UN COMTRADE, then classified and recalculated to form the following variable: Animal husbandry export of each country selected for the analysis to other EU partners (27 in total). Re-exports were not considered due to unavailable data. Since these were given in current prices, to avoid evident distortion of the results a decision was made to apply price indices (real price adjusted indices of agricultural products, output, annual data, 2010=100) and use values in constant prices for further analysis. Price indices for each country were taken from Eurostat.

Due to a substantial difference in the extent to which the countries selected for the analysis possess agricultural land (and correspondingly in volumes of their export) a decision was made to recalculate the volume of export by each country per hectare of agricultural land available in the country. Considering that it seemed to serve no justified purpose to encumber the analysis with additional data on arable land, these were eventually not considered.

Similarly, a decision was made to recalculate export volumes per person employed in agriculture and per million USD of Fixed Capital Consumption (Agriculture, Forestry and Fishing). The data on Employment in agriculture and Consumption of Fixed Capital were retrieved from FAOSTAT.

All the data available and derived this way, along with constructed variables, were compared and balanced to avoid lacking observations and to represent the same time span for each country. Regrettably, Hungary and Italy were excluded from the analysis since no data on price indices for the period from 2005/2007 to 2010 were available. As a result, a consistent dataset of 104 observations in total, representing Austria, Belgium, Czechia, Denmark, France, the Netherlands, Poland, and Slovakia, incorporates the period from 2005 to 2017. The data for Germany were not included as these refer to a shorter period from 2010 to 2017 (again, the problem concerned unavailability of the data on price indices).

The Table 1 provides a summary of the data used in the analysis.

The export and import data were retrieved from the UN COMTRADE database as mentioned above. They are based on the Standard International Trade Classification (SITC) Rev. 2 nomenclature. A one-digit level of aggregation was used, consisting of the products mentioned in Table 2.

The initial analysis was conducted at a one-digit level of aggregation. In this regard, only four commodity aggregates that include animal husbandry were analysed in more detail. This concerns group 00, 01, 02 and 41 (Table 3).

Variable	Label	Units of measurement
Animal husbandry export	aEx	USD, constant 2010 prices
Agricultural land	land	1000 ha
Employment in agriculture	labour	1000 persons
Consumption of Fixed Capital (Agriculture, Forestry and Fishing)	capital	USD, constant 2010 prices
Animal husbandry export per ha	aEx.p.ha	USD, constant 2010 prices /ha
Animal husbandry export per person empl.	aEx.p.worker	USD, constant 2010 prices/worker
Animal husbandry export per $1 million of fixed capital consumption	aEx.p.capital	USD, constant 2010 prices

Source: own elaboration.

Table 1: Summary of created variables.

Variable	Obs	Mean	Std.Dev.	Min	Max
land	104	7300.524	9231.325	1327	29390.4
labour	104	448.14	640.76	53.02	2452.089
capital	104	3842.832	4508.11	408.88	15674.72
Variable	Obs	Mean	Std.Dev.	Min	Max
aEx	104	6.30e+09	5.01e+09	3.35e+08	1.88e+10
aExpha	104	2380.042	2999.8	131.374	10208.41
aExpworker	104	42377.73	47136.16	1046.999	163000
aExpcapital	104	2417814.60	2099174.77	624872.40	8572369.4

Source: authors' elaboration in STATA

Table 2: Descriptive Statistics of variables.

Code	Description
00	Live animals other than animals of division 03
01	Meat and meat preparations
02	Dairy products and birds' eggs
41	Animal oils and fats

Source: authors' elaboration in STATA

Table 3: Commodity groups in analysis.

Years 2005, 2010 and 2017 were selected for the analysis, since 2005 is the very first year of the Czech membership in the EU, 2010 represents the year after the financial crisis, and 2017 includes the latest full dataset available for productivity analysis.

Research methods

The objective of this research is to analyse comparative advantages at the state and European level with regard to the productivity measures. There are different assessment indicators for this comparison which were used to explore foreign trade of the monitored countries.

Firstly, traditional competitiveness indexes were calculated. These include the Balassa index that was calculated according to Laursen (2015) with the threshold effect equalling 1. RCA >1 – the country possesses a competitive advantage. RCA < the country possesses a competitive disadvantage. The RCA index was calculated in two ways. The bilateral RCA was calculated in which the situation between the Czech Republic and other countries is assessed (the Czech Republic is the benchmark) and then the position of the countries against the EU level (RCA EU).

To evaluate the intra-industry trade the Gruber Lloyd Index (GLI) (Grubel and Lloyd, 1971) was used. GLI = 1 – only intra-industry trade exists GLI = 0 – there is no intra-industry trade, only inter-industry trade. Finally, the Lafay index (LFI) was employed to assess mutual trade (Iapadre, 2001; Lafay, 1992). It attains values <-∞,∞>, if the value exceeds zero, the country possesses a comparative advantage.

Based on the above indicators, cluster analysis was conducted, in which individual calculation was used as an input variable for cluster analysis. Hierarchical clustering and Ward´s method (Ward, 1963) were used. To prevent distortion the variables were transformed using the z-score. The final step was to assess the differences between the groups.

Multidimensional scaling was used and perceptual maps were created for graphic illustration (Buja and Swayne, 2002; Torgerson, 1952).

Results and discussion

Export-measured productivity of selected agricultural sectors in EU

To analyse productivity of agricultural, specifically animal husbandry in the European countries selected for the analysis (based on available balanced data incorporating the period from 2005 to 2017), a decision was made to trace export performance of these countries recalculated on a unit of core productive factors, such as land, labour and capital. The figures below provide graphic representation of the mentioned indicators' development in time for all the countries.

Since the collected export data were initially given in current prices, to avoid evident distortion of comparison results, corresponding price indices were applied to time series of each country to express all export data in constant prices (as per 2010 year)[1].

Figure 1 illustrates the dynamics of export values in selected European countries, where individual year-to-year values are given in constant as of 2010 prices, bln. USD[2].

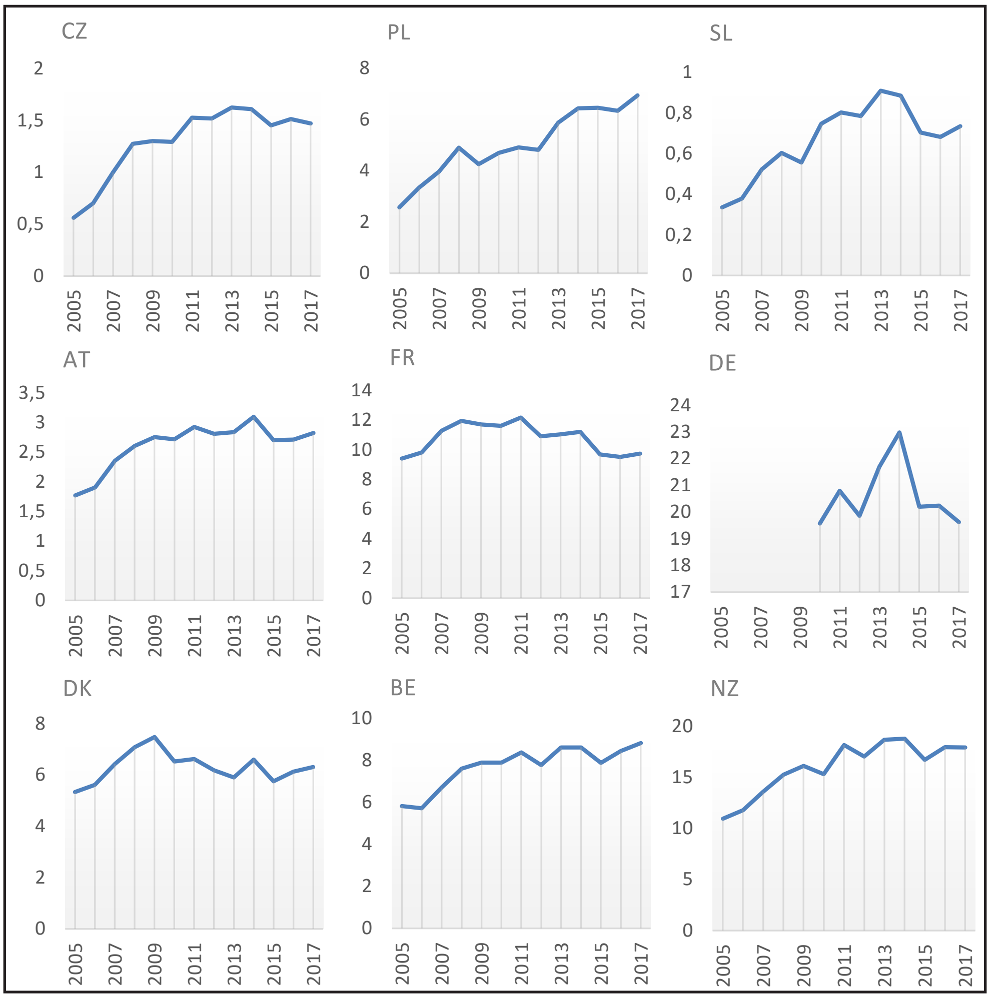
It is obvious that the values provided above cannot be compared directly (due to the countries' sizes, population densities and corresponding sizes of their agricultural lands), these numbers will be recalculated with regard to a unit of core productive factors: per hectare of agricultural land, per worker employed in the agricultural sector and per million USD of fixed capital consumption.

As a result, the next Figure 2 shows the dynamics of export values in selected European countries per hectare of agricultural land disposable for the country in question. All values are also given in constant 2010 prices, USD.

As can be seen from the Figure 2, in terms of capability of benefiting from their agricultural land (export-measured productivity), the leaders are the Netherlands, Belgium and Denmark, followed by Germany, Austria, France, Slovakia, Poland and Czechia (ranked from highest to lowest).

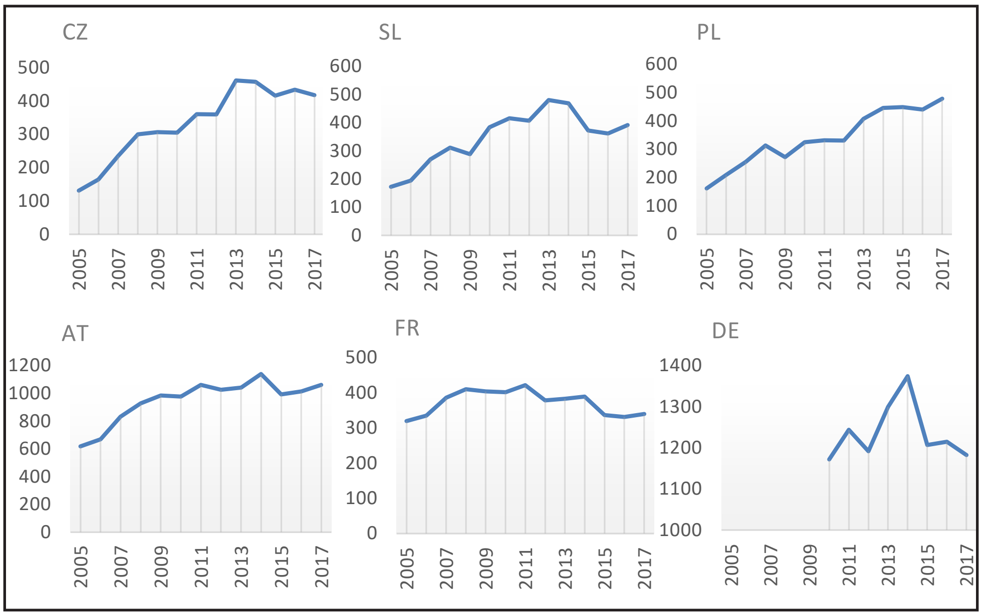
1 Price indices were retrived from Eurostat database [5.06.2020]

2 https://ec.europa.eu/eurostat/data/database?p_p_id=NavTreeportletprod_WAR_NavTreeportletprod_INSTANCE_nPqeVbPXRmWQ&p_p_lifecycle=0&p_p_state=normal&p_p_mode=view&p_p_col_id=column-2&p_p_col_pos=2&p_p_col_count=3



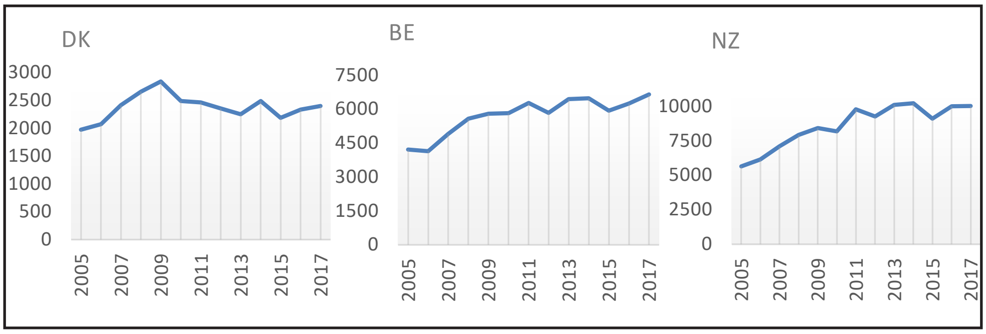
Source: authors' elaboration

Figure 1: Animal husbandry exports, in bln. USD (constant 2010 prices).



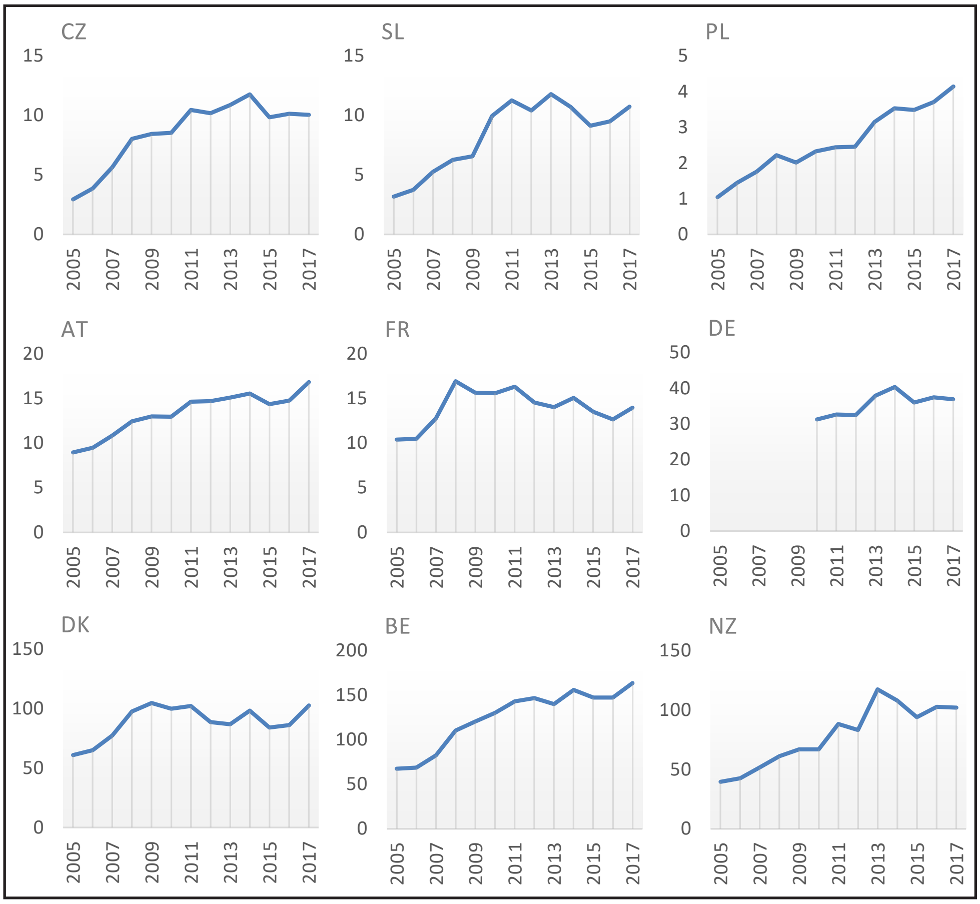
Source: authors' elaboration

Figure 2: Animal husbandry export per 1 ha, USD (constant 2010 prices) (to be continued)



Source: authors' elaboration

Figure 2: Animal husbandry export per 1 ha, USD (constant 2010 prices) (continuation).



Source: authors' elaboration

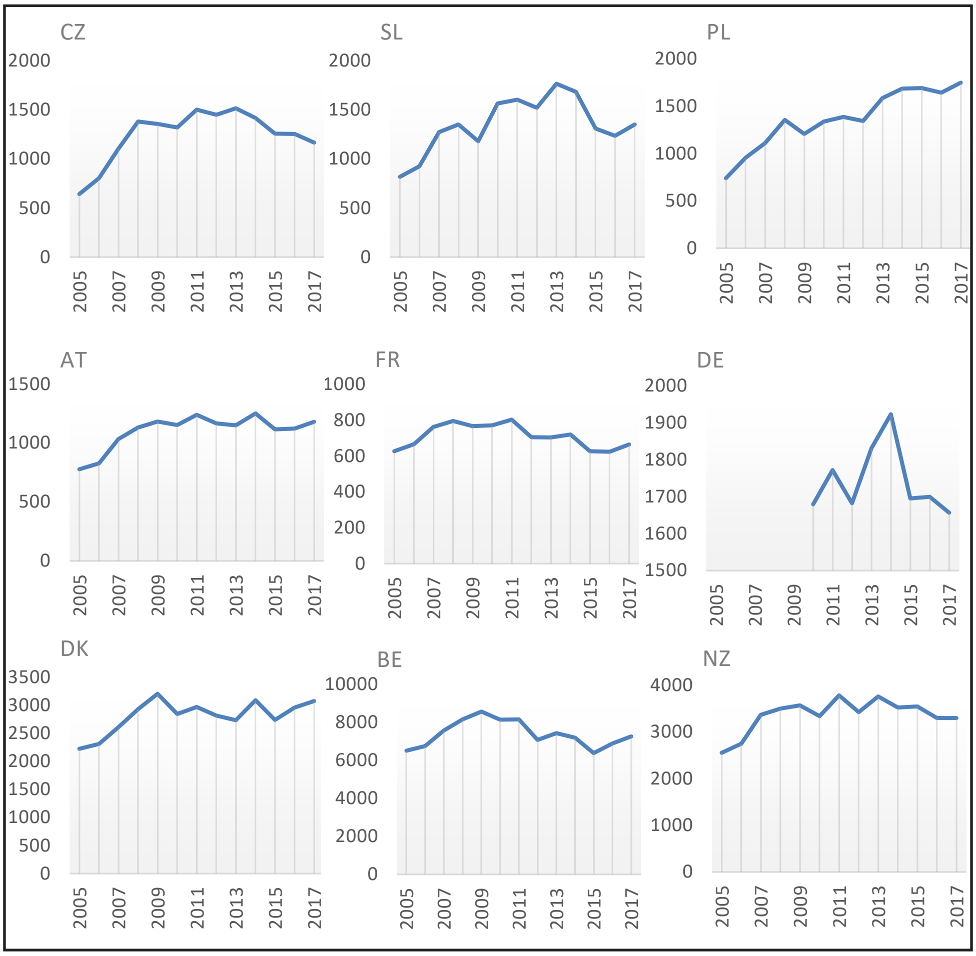
Figure 3: Animal husbandry export per 1 worker employed, thousands USD (constant 2010 prices).

Regarding the values of Animal husbandry export per worker employed, the situation is slightly different, although the overall composition of leaders is almost identical: Belgium, Denmark and the Netherlands are followed by Germany, France, Austria, Czechia, Slovakia, and Poland. Figure 3 provides graphs corresponding to each country.

When considering the countries' capability of benefiting from their fixed capital available for the agricultural sector, Belgium is the evident leader among the countries selected for the analysis. Approximately twice as low values of export per million USD of fixed capital consumption are recorded the Netherlands and Denmark, followed by Germany. During the analysed period, Poland and Slovakia displayed almost identical values, whereas Czechia, which surpassed Austria, attained export values per million USD of fixed capital consumption 1.75 times higher than France.

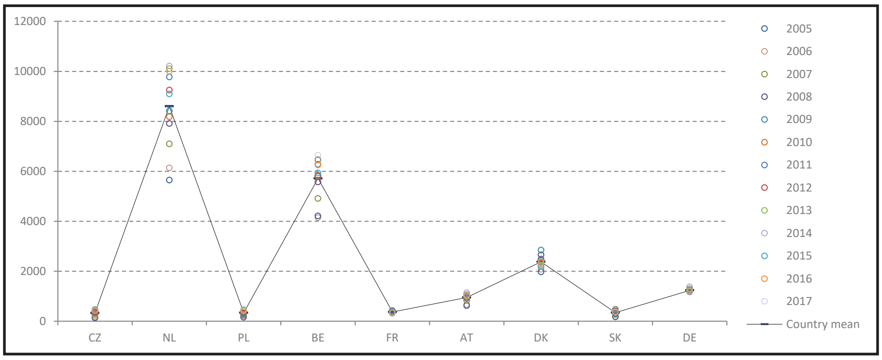
The corresponding graphs are provided in Figure 4.

Simultaneously, in order to clearly illustrate the borders within which the observed values of exports per unit of core productive factors fluctuate along with inter-country annual mean values and year-on-year mean values per country, the following figures are provided: Figure 5 which shows the comparison of mean values of each country's export per hectare for the entire analysed period from 2005 to 2017, Figure 6 in turn illustrates the comparison of export values per hectare among all 8 countries for each year along with the inter-country annual mean values.



Source: authors' elaboration

Figure 4: Animal husbandry export per $1 million of fixed capital consumption, thousands USD (constant 2010 prices).



Note: In case of Germany the analysis covers the period from 2010-2017

Source: authors' elaboration.

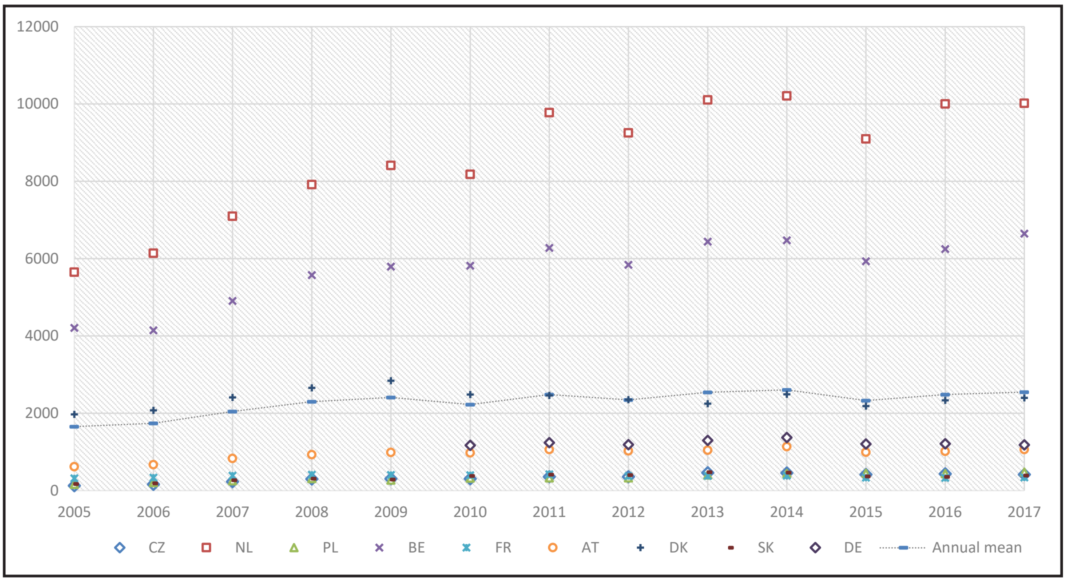
Figure 5: Animal husbandry export per 1 ha and countries' mean values calculated for the period from 2005 to 2017, in USD (const. 2010 prices).

As can be seen from the Figure 5, there is a considerable gap between the countries' mean values, which indeed suggests the very existence of space to improve export-measured productivity for countries such as Czechia, Poland and Slovakia, as the lowest mean values of export per hectare were recorded in these countries. France, Austria, Germany and Denmark can potentially increase their productivity as well.

However, it is worth mentioning that in contrast to the absolute values of exports per unit of core production factors, the highest on average year-on-year growth rates in exports within the analysed period were recorded in Slovakia, Czechia and Poland. The summary Table 4 below provides the average year-on-year growth rates in exports per hectare of agricultural land, per worker employed in the agricultural sector and per million USD of fixed capital consumption.

Considering the above values of year-on-year growth rates, it is possible to assert that Slovakia, Czechia and Poland are the countries with the highest potential to achieve better export-measured productivity in the following years, they appear to be in pursuit of enhancing their export productivity per unit of core production factors.

The most considerable decline in export volumes



Note: In case of Germany the analysis covers the period from 2010-2017
Source: authors' elaboration.

Figure 6: Animal husbandry export per 1 ha and inter-country annual mean values, in USD (const. 2010 prices).

Average growth rate in exports per 1 ha of agricultural land		Average growth rate in exports per 1 worker employed		Average growth rate in exports per $1 million of fixed capital consumption	
13.56	**SK**	17.71	**SK**	9.76	**SK**
13.23	**CZ**	15.17	**CZ**	8.65	**CZ**
9.44	**PL**	11.93	**PL**	7.41	**PL**
5.44	AT	8.33	BE	4.28	AT
5.08	NL	8.17	NL	2.86	DK
4.18	BE	5.11	AT	2.48	NL
1.88	DK	4.74	DK	1.60	BE
0.70	FR	2.98	FR	0.70	FR
0.20	DE	1.44	DE	0.02	DE

Note: In case of Germany the analysis covers the period from 2010-2017
Source: authors' elaboration.

Table 4: The average values of year-on-year growth rates in Exports per a unit of core productive factors for the period from 2005 to 20171), in %.

during the analysed period was registered in almost all the countries in the same year, 2015, when the Russian embargo was imposed on the European partners as a direct response to the European sanctions against Russia due to different attitudes to the Crimean events. Table 5 summarizes the most significant negative year-on-year growth rates in the countries' exports observed within the period from 2005 to 2017 and a particular year when this decline occurred.

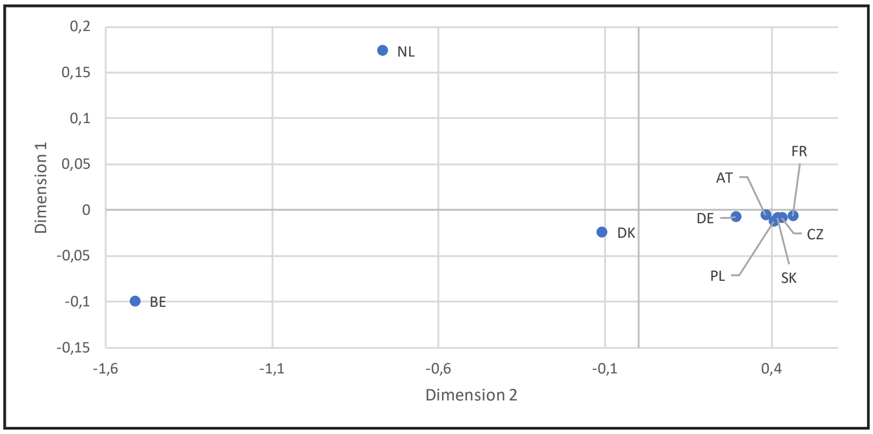
Multidimensional scaling was applied to the above productivity indicators connected to foreign trade to determine the main groups of countries (Figure 7).

Based on the multidimensional scaling, the countries can be divided into two groups. The descriptive analysis of both is included in Table 6. The first group, which consists of Belgium, the Netherlands and Denmark,

Maximal negative growth rates in exports per 1 ha of agricultural land			Maximal negative growth rates in exports per 1 worker employed			Maximal negative growth rates in exports per $1 million of fixed capital consumption		
SK	-20.44	**2015**	CZ	-16.40	**2015**	SK	-22.17	**2015**
FR	-13.40	**2015**	SK	-14.70	**2015**	BE	-13.37	2012
PL	-13.30	2009	DK	-14.42	**2015**	FR	-12.96	**2015**
AT	-12.81	**2015**	NL	-12.97	**2015**	DE	-11.85	**2015**
DK	-12.48	2010	FR	-10.78	2012	DK	-11.41	**2015**
DE	-12.15	**2015**	DE	-10.70	**2015**	CZ	-11.22	**2015**
NL	-10.90	**2015**	PL	-9.11	2009	PL	-11.00	2009
CZ	-9.24	**2015**	AT	-7.58	**2015**	AT	-10.75	**2015**
BE	-8.39	**2015**	BE	-5.56	**2015**	NL	-9.48	2012

Note: In case of Germany the analysis covers the period from 2010-2017
Source: authors' elaboration.

Table 5: The biggest negative year-on-year growth rates in exports per a unit of core productive factors registered in the period from 2005 to 20171), in %.



Source: authors' elaboration

Figure 7: Multidimensional scaling of productivity indicators related to foreign trade (2017).

	aEx.p.ha_const.pr.	aEx.p.worker_const.pr.	aEx.p.capit._const.pr.
Mean group 1 (NL, BE, DK)	6 353.38	122 579.81	4 544 669.05
Mean group 2	645.56	15 436.59	1 295 493.34
Std. Deviation group 1	3 815.64	35 374.41	2 356 148.20
Std. Deviation group 2	374.15	11 335.66	392 156.47

Source: authors' elaboration.

Table 6: Descriptive analysis of the productivity indicators related to foreign trade (2017).

is rather heterogenous with regards to their productivity. Belgium attains the highest productivity level per worker and capital. The average productivity in this group is multiple times higher than in the second. The standard deviation is rather high, maximum for capital and the lowest for land. Therefore, the means of the indicators in this group differ significantly from those in the second group, which are very similar and concern Austria, the Czech Republic, France, Germany, Poland, and Slovakia.

When comparing the situation in 2005 and 2010, only slight differences are detected between the years. There is an increase in standard deviation and the mean of these variables. In relative numbers, the highest increase in standard deviation is for labour productivity, followed by land and the lowest relative increase is for capital.

When the countries are divided into "old" and "new" member states, what can be observed is that the highest relative increase of standard deviation for the "new" member states concerns productivity of capital, which is followed by labour. In the case of the "old" member states, it is the capital followed by land. It could thus be concluded that there is an important increase in the productivity of capital and labour in the new member states.

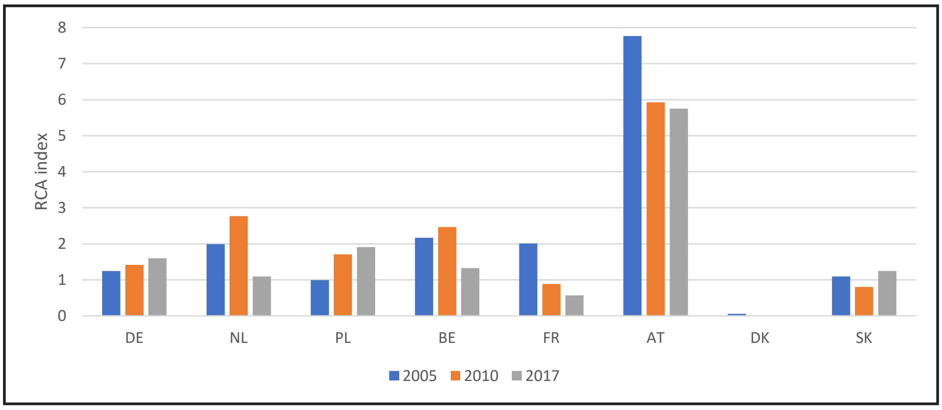
Indicators of competitiveness

RCA – bilateral – the relationship to the Czech Republic (comparative advantage of the Czech Republic over the examined countries)

Figure 8 displays the comparison of the RCA index for live animals. It is evident that the Czech Republic possessed a comparative advantage over Austria in 2017; however, compared to 2005, it witnessed a slight decrease. The results also show a strong position of the Czech Republic against the Netherlands, Belgium or Poland. A very high RCA over Austria relates to the structure of the production since Czechia exports enormously high volumes of live animals to Austria where slaughterhouses with higher redemption price than in Czechia are located. When comparing the situation between 2005 and 2017, Czechia increased its comparative advantage in three states (DE, PL, SK) and decreased its position in 5 cases (NL, BE, FR, AT, DK). This might be alarming for the future.

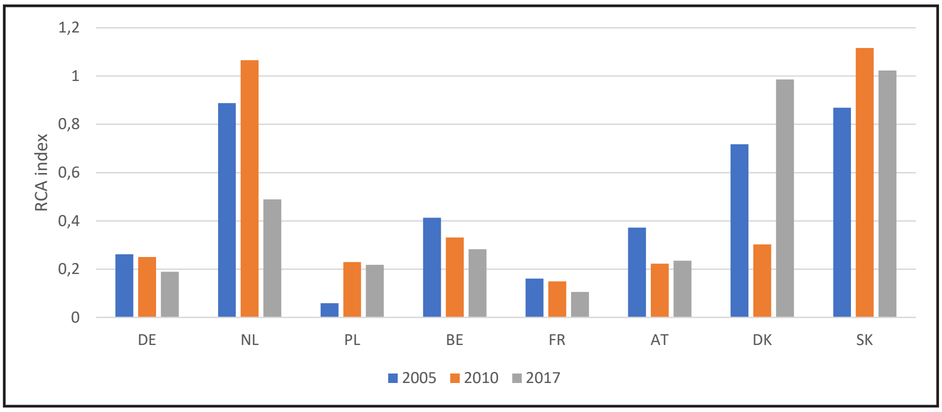
The situation is similar in the case of meat and meat preparations. The Czech Republic has decreased its comparative advantage over 7 countries and increased only over Denmark and Slovakia. The situation is relatively stable for the future long-term period and the RCA is relatively high. The value of RCA is lower, and the distribution is more homogenous than in the case of live animals. However, the Czech Republic's RCA > 1 only over Slovakia and reaches approximately the same value as Denmark. Regarding other states, the Czech Republic does not possess a comparative advantage in meat and meat preparations.

The number of Czechia´s comparative advantages of dairy products and eggs has also declined over time. In 2017, the Czech Republic possessed a comparative advantage only over two countries – Germany and Slovakia, while in 2005 it was four. The country has lost its position against the Netherlands or Denmark, which is rather alarming.



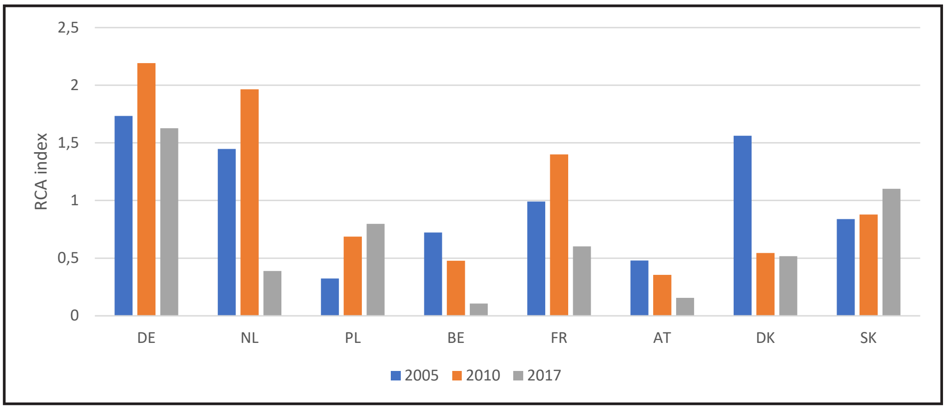
Source: authors' elaboration

Figure 8: RCA index live animals – comparison (2005, 2010, 2017).



Source: authors' elaboration.

Figure 9: RCA index Meat and meat preparations – comparison (2005, 2010, 2017).



Source: authors' elaboration

Figure 10: RCA index Dairy products and birds' eggs (2005, 2010, 2017).

RCA – relationship at the European level

Figure 11 displays the distribution of comparative advantages of the selected countries on the European market. The median value is quite similar for all the analysed years and commodity groups; however, the distribution changed during the monitored years. In general, it can be stated that the smallest differences are between the commodity aggregate dairy products and birds' eggs (S3-02). The comparative advantage of all the European countries is rather similar. The most significant changes have been monitored for the commodity aggregate live animals (S3-00).

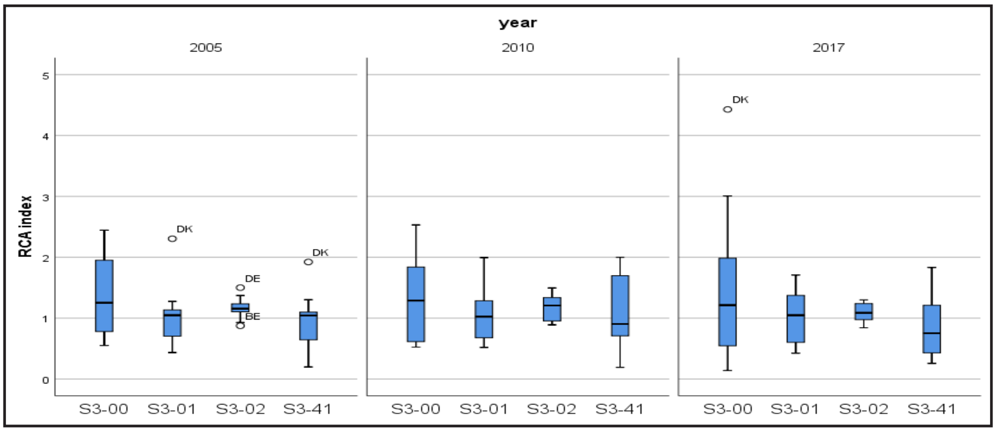
Correlation analysis was applied to the RCA indexes (European level) and export measured productivity indicators. This analysis includes CZ, DE, NL, PL, BE, FR, AT, DK, and SK.

A correlation exists between RCA indexes (European level) and export measured productivity indicators. There is a negative correlation between live animals (S3-00) and all productivity indicators, although only land is statistically significant. A similar situation exists between dairy products and birds' eggs. In this case the only significant correlation is between capital and RCA. The most considerable significance is between RCA of animal oils and fats and labour productivity.

What could be stated is that the factor of productivity does not influence the comparative advantage of these states.

After assessing the comparative advantage of the Czech Republic over the European Union during the monitored years it might be concluded that there is a slight decrease in live animals and dairy products and birds 'eggs. These two commodity aggregates possess a comparative advantage on the European market. Regarding meat and meat preparations and animal oils and fats, the Czech Republic does not possess a comparative advantage.

When comparing the remaining European countries, it is evident that Denmark possesses the most significant comparative advantage in the case of live animals.



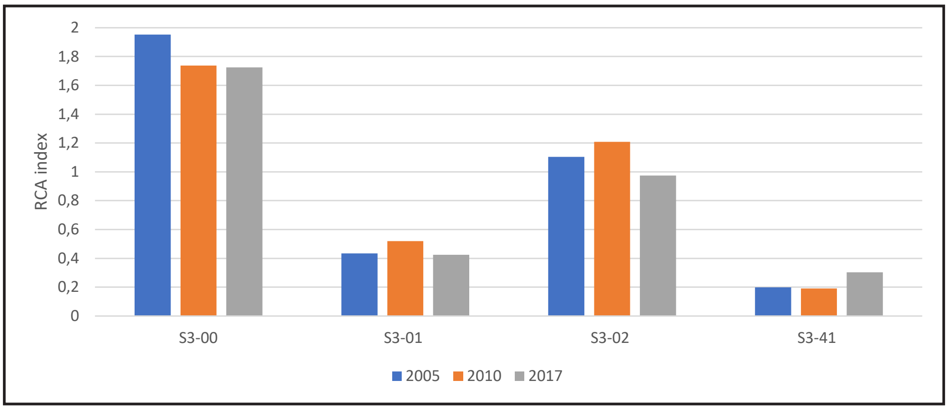
Source: authors' elaboration

Figure 11: Distribution of RCA index with regards to the European level.

	aEx.p.ha_const.pr.	aEx.p.worker_const.pr.	aEx.p.capit._const.pr.
RCA S3 - 00	-.386*	-.160	-.342
RCA S3 - 01	.391*	.256	.374
RCA S3 - 02	-.248	-.162	-.395*
RCA S3 - 41	.299	.490**	.223

Note: * Correlation is significant at the 0.05 level (2-tailed); ** Correlation is significant at the 0.01 level (2-tailed)
Source: authors' elaboration

Table 7: Correlation between RCA index and export measured productivity indicators.



Source: authors' elaboration

Figure 12: Evolution of RCA´s of the Czech Republic x EU.

Division of the countries according to export-based productivity measures

Before applying the clustering procedure, correlation analysis was applied to determine the connection between the original variables. Based on this, the LFI indicator was excluded due to a high correlation between the RCA and LFI. Based on the hierarchical clustering technique and Ward's method, similarities between the monitored countries were identified.

The commodity aggregate S3-00 - Live animals (Tables 8, 9, 10) does not witness any significant changes in the monitored period. The first group of countries consists of France, Slovakia, Germany, and Poland. The group mostly possesses a comparative advantage in this commodity group. The productivity connected with land is lower than in the second and third group. The standard deviation for productivity measures with regard to land is lower. The highest productivity is recorded in Germany.

The situation in the second group, which includes Austria only, has already been mentioned above. The reason why Austria is separated is a very high

DESCRIPTION	GROUP 1	GROUP 2	GROUP 3	GROUP 4
00 Live animals other than animals of division 03	FR, SK, DE, PL	AT	NL, BE, DK	
01 Meat and meat preparations	PL, FR, DE, AT	DK, SK	NL, BE	
02 Dairy products and birds' eggs	PL, SK, FR	AT, DK	DE	NL, BE
41 Animal oils and fats	PL, FR, AT	DE, SK	DK	NL, BE

Source: authors' elaboration

Table 8: Division of countries (2017).

DESCRIPTION	GROUP 1	GROUP 2	GROUP 3	GROUP 4
00 Live animals other than animals of division 03	DE, SK, FR, PL	AT	NL, BE, DK	
01 Meat and meat preparations	DE, PL, AT, FR, DK	BE	NL, SK	
02 Dairy products and birds' eggs	AT, SK, PL, DK	BE	DE, FR, NL	NL, BE
41 Animal oils and fats	PL, FR, DE, AT	SK	NL, DK, BE	NL, BE

Source: authors' elaboration

Table 9: Division of countries (2010).

DESCRIPTION	GROUP 1	GROUP 2	GROUP 3	GROUP 4
00 Live animals other than animals of division 03	PL, SK, FR, DE	AT	BE, DK, NL	
01 Meat and meat preparations	PL, FR, DE, AT	DK	NL, BE, SK	
02 Dairy products and birds' eggs	DE, DK, NL, BE	FR, SK, AT, PL		NL, BE
41 Animal oils and fats	DE, FR, AT, PL,	SK	NL, DK, BE	NL, BE

Source: authors' elaboration

Table 10: Division of countries (2005).

level of the RCA index as well as the existence of a comparative advantage. The last group of countries includes the Netherlands, Belgium and Denmark. All of these reach a high productivity level with regard to factors of production.

The second commodity aggregate includes meat and meat preparations. The division of groups changed in the monitored years; the first group contains Poland, France, Austria and Germany, while in 2010 it also included Denmark. The median value of productivity of land is the lowest of all the monitored countries and the standard deviation is also relatively low, thus this is a relatively homogenous group. Moreover, productivity per worker is the lowest of all the monitored groups with the lowest standard deviation. This group has the lowest median value of RCA with very low dispersion. This group might be referred to as the countries possessing a relative comparative disadvantage over the Czech Republic with weak export-based production performance.

The second group of countries (Denmark and Slovakia) is rather heterogenous with regard to export-based production performance, except for land. The RCA is the highest of all the monitored groups. The third group (the Netherlands and Belgium) records very strong export-based production performance gaining a comparative disadvantage only over the Czech Republic. These two groups changed during the analysed years.

Dairy products and birds' eggs (S3-02) and animal oils and fats (S3-41) have developed as the most diverse commodity aggregates. In 2017, they were divided into 4 sub-groups.

In the case of dairy products and birds' eggs, there were only two groups in 2005, three in 2010, and four in 2017. This indicates that that the export-oriented production performance and indicators of competitiveness changed in the monitored years. The first group includes Poland, Slovakia and France and has the lowest mean for productivity indicators; however, it also has a relatively high RCA (the highest value of all the selected groups). The mean value of GLI is approximately 0.6. All the indicators have the lowest standard deviation. The second group includes Denmark and Austria which have the highest GLI with a very high mean value of labour productivity. The last group includes the Netherlands and Belgium and has the strongest export-based production performance with the lowest GLI and bilateral RCA over Czechia.

The final group of products is specific in that its production fluctuates enormously between the years.

Discussion

The existing literature express a clear connection between international trade (especially exports and factor productivity and its growth (Bhagwati, 1978). However, the theoretical and empirical literature focusing on agricultural factor productivity related to foreign trade is rather limited (Sunge and Ngepah, 2020).

The presented results clearly show that in the monitored group of EU countries there is a within-country difference in factor export-measured productivity indicators as well as in revealed comparative advantages. However, it cannot be concluded that these differences in animal production export measures differ only between the old and the new member states. There are countries like France, Austria, or Germany where factor productivity is more similar to the new member states. On the other hand, the new member states have witnessed a considerably higher average growth rate in export per hectare of agricultural land, per worker employed or per fixed capital consumption. Based on the average growth rate three main groups can be identified. The first includes the new member states with the highest growth rate (SK, CZ, PL), followed by middle growth countries NL, BE, AU, and DK. France and Germany record the lowest growth rate. The findings support the idea of Kijek at al. (2019) about convergence in agriculture and lower productivity growth of Germany and France. It is quite interesting because when the export-based productivity performance is analysed, these two countries are more similar to the new member states than to the Netherlands, Belgium, and Denmark.

The productivity of animal husbandry differs across the monitored countries. The reason why it can vary significantly includes technology transfers, resource allocation, competition, or the use of economies of scale. On the other hand, gains resulting from international trade are connected with the existence of comparative advantages and the utilisation of economies of scale and thus increasing return to scale or openness of economy. Contrary, (Ciaian et al., 2009 have discovered that the revealed comparative advantage does not depend solely on economies of scale but also on the type of the farm, since family farms focus more on labour intensive products and can have comparative advantages compare to corporate farms which are more capital intensive. It also supports the finding about the situation in the Czech Republic that there is a greater concentration of animal producers.

Consequently, factor productivity influences effective results of foreign trade and vice versa. However, Tong Soo (2013) argues that the gain for small countries is always more considerable than in large countries. This would mean that the Czech Republic should be able to use its productivity better than for example Poland. When comparing the position of the Czech Republic and Poland in terms of their comparative advantage, the findings indicate that in the case of their bilateral agreement there is a comparative advantage with regard to live animals over the majority of the analysed countries except Denmark and France. The problem of the Czech agricultural production is deterioration of its position in the case of Meat and meat preparations and Dairy products and birds' eggs at the European level. In comparison with live animals, where the latter records a revealed comparative advantage, the country has lost its comparative advantage which has become a disadvantage. These two commodity aggregates consist of products with a slightly higher value added than live animals only. It might be stated that the situation with the comparative disadvantage is more stable and it does not fluctuate as much as the comparative advantage. The same has been observed by (Qineti, Rajcaniova and Matejkova, 2009) in the case of Slovakia.

(Abizadeh and Pandey, 2009) have discovered that trade openness does not have a positive effect on factor productivity in agriculture, although it has a positive impact on an entire national economy.

However, one of the factors that influence the position of the country on the international market is the existence of retail companies which might have both positive and negative impact on the overall competitiveness of the country on the international market. The question is whether it is, in fact, the factor productivity, economies of scale and specialization that affect competitiveness of countries on international markets or whether there are other business powers that might influence the situation of agricultural sectors.

Conclusion

In this paper data for 9 EU member states have been used to investigate the relationship between factor export measured productivity indicators and the revealed comparative advantage for animal husbandry. Years 2005, 2010 and 2017 have been compared and an existing gap in literature has been highlighted, which implies that it does not focus on the connection between these indicators and comparative advantages together with competitiveness.

The evidence from this study suggests that the revealed comparative advantage of these countries is not determined primarily by the level of export measured productivity. The relationship between these variables is rather weak and very often negative. This means that productivity indicators do not play a significant role in the overall competitiveness of the monitored countries. The sectors of animal husbandry in which the Czech Republic has a comparative advantage have been identified and the fact that, concerning the production itself, Czechia focuses more on products without higher added value (life animals) has been emphasised.

When the countries are divided according to their export-based productivity performance it might be stated that there are similarities between them. It is rather surprising that France and Poland are indicated in one group (with the Czech Republic as the benchmark), very often accompanied by Slovakia, Germany, and Austria. In contrast, Belgium and the Netherlands are also in the same position. Based on the above results, it might be concluded that there are differences between the countries and that the Czech Republic has a unique position with regard to export-based productivity performance.

Acknowledgements

The research results are part of the project QK1920398 "Duality in Czech Agriculture: Advantage or Disadvantage for New Generation Agriculture?" within the Program of Applied Research "ZEMĚ" of the Ministry of Agriculture for 2017-2025 funded by the Ministry of Agriculture of the Czech Republic, NAZV.

Corresponding authors
prof. Ing. Luboš Smutka, Ph.D.
Department of Economics, Faculty of Economics and Management
Czech University of Life Sciences Prague, Kamýcká 129, Prague - Suchdol, 165 00, Czech Republic
Phone: +420224382076, E-mail: smutka@pef.czu.cz

References

[1] Abizadeh, S. and Pandey, M. (2009) "Trade Openness, Structural Change and Total Factor Productivity", *International Economic Journal*, Vol. 23, No. 4, pp. 545–559. E-ISSN 1743-517X, ISSN 1016-8737. DOI 10.1080/10168730903372273.

[2] Berthou, A., Dhyne, E., Bugamelli, M., Cazacu, A, Demian, V., Harasztosi, P., Lalinsky, T. Meriküll, J., Oropallo, F. and Soares, A. C. (2015) "Assessing European Firms' Exports and Productivity Distributions: The CompNet Trade Module", ECB Working Paper No. 1788. [Online]. Available: https://ssrn.com/abstract=2621440. [Accessed: Oct. 20, 2020].

[3] Bojnec, Š. and Fertö, I. (2014) "Meat Export Competitiveness of European Union on Global Markets", *Agricultural and Food Science,* Vol. 23, No. 3. ISSN 1795-1895. DOI 10.23986/afsci.9373.

[4] Bojnec, Š. and Fertö, I. (2015) "Agri-food export Competitiveness in European Union Countries", *Journal of Common Market Studies*, Vol. 53, No. 3, pp. 476-492. ISSN 1468-5965. DOI 10.1111/jcms.12215.

[5] Buja, A. and Swayne, D. F. (2002) "Visualization Methodology for Multidimensional Scaling", *Journal of Classification*, Vol. 19, No. 1, pp. 7-43. E-ISSN 1432-1343, ISSN 0176-4268. DOI 10.1007/s00357-001-0031-0.

[6] Burianová, J. (2010) "Trends of the Agrarian Foreign Trade of CR after Accession to EU, Competitiveness of Commodities", *AGRIS on-line Papers in Economics and Informatics,* Vol. 12, No. 3. ISSN 1804-1930.

[7] Carraresi, L. and Banterle, A. (2008) "Measuring Competitiveness in the EU market: a Comparison between Food Industry and Agriculture", *Proceedings of EAAE 2008 International Congress*, Ghent. DOI 10.22004/ag.econ.43692.

[8] Ciaian, P., Pokrivcak, J. and Drabik, D. (2009) "Post-Communist Economies Transaction costs, product specialisation and farm structure in Central and Eastern Europe", *Post-Communist Economies*, Vol. 21, No. 2, pp. 191-201. E-ISSN 1465-3958, ISSN 1463-1377. DOI 10.1080/14631370902778526.

[9] Csikósová, A., Čulková, K. and Janošová, M. (2018) "Work Productivity and Employment in chosen EU Countries", *Proceedings of 8th Business and Managements Conference*, Venice. ISBN 978-80-87927-73-1. DOI 10.20472/BMC.2018.008.002.

[10] Čechura, L., Žáková Kroupová Z. and Hockmann, H. (2015) "Market Power in the European Dairy Industry", *AGRIS on-line Papers in Economics and Informatic*s, Vol. 7, No. 4, pp. 39-47. ISSN 1804-1930. DOI 10.7160/aol.2015.070404.

[11] Fertö, I. and Hubbard, L. J. (2003) "Revealed Comparative Advantage and competitiveness in Hungarian Agri-food Sectors", *The World Economy*, Vol. 26, No. 2. ISSN 1467-9701. DOI 10.1111/1467-9701.00520.

[12] Fojtíková, L. and Staníčková, M. (2017) "Konkurenceschopnost a produktivita vývozu členských zemí Evropské unie", *Politická ekonomie*, Vol. 65, No. 6, pp. 669-689. ISSN 2336-8225. DOI 10.18267/j.polek.1169. (in Czech).

[13] Galović, T., Mišević, P. and Popović, L.(2017) "The Export Competitiveness of the EU 28 Countries", *Trade Perspectives 2017*, University of Zagreb - Faculty of Economics and Business, Croatian Chamber of Economy. ISBN 978-953-346-048-2.

[14] Gorton, M., Davidova, S. and Ratinger, T. (2000) "The Competitiveness of Agriculture in Bulgaria and the Czech Republic Vis-à-Vis the European Union (CEEC and EU Agricultural Competitiveness)", *Comparative Economic Studies*, Vol. 42, No. 1. ISSN 1478-3320. DOI 10.1057/ces.2000.2.

[15] Grubel, H. G. and Lloyd, P. J. (1971) "The Empirical Measurement of Intra-Industry Trade", *Economic Record*, Vol. 47, No. 4, pp. 494-517. E-ISSN 1475-4932. DOI 10.1111/j.1475-4932.1971.tb00772.x.

[16] Iapadre, P. L. (2001) "Measuring international specialization", *International Advances in Economic Research*, Vol. 7, No. 2, pp. 173-183. E-ISSN 1573-966X, ISSN 1083-0898. DOI 10.1007/BF02296007.

[17] Kharin, S. (2019) "Horizontal Price Transmission on the Russian Dairy Market: Nonlinear Approach", *AGRIS on-line Papers in Economics and Informatics*, Vol. 11, No. 3, pp. 45-54. ISSN 1804-1930. DOI 10.7160/aol.2019.110305.

[18] Kijek, A., Kijek, T., Nowak, A. and Skrzypek, A. (2019) "Productivity and its convergence in agriculture in new and old European Union member states", *Agricultural Economics*, Vol. 65, No. 1, Vol. 01-09. E-ISSN 1805-9295, ISSN 0139-570X. DOI 10.17221/262/2017-AGRICECON.

[19] Lafay, J. (1992) "The Measurement of Revealed Comparative Advantages", In M. G. Dagenaisand, P. A. Muet (Eds.) "*International Trade Modeling*", pp. 209-234, London, United Kingdom: Chapman and Hall.

[20] Latruffe, L. (2010) "Competitiveness, Productivity and Efficiency in the Agricultural and Agri-Food Sectors", *OECD Food, Agriculture and Fisheries Papers*, No. 30, OECD Publishing, Paris. E-ISSN 18156797. DOI 10.1787/5km91nkdt6d6-en.

[21] Laursen, K. (2015) "Revealed comparative advantage and the alternatives as measures of international specialization", *Eurasian Bussines Review*, Vol. 5, No. 1, pp. 99-115. E-ISSN 2147-4281. DOI 10.1007/s40821-015-0017-1.

[22] Mizik, T., Gál, P. and Török, Á. (2020) "Does Agricultural Trade Competitiveness Matter? The Case of the CIS Countries", *AGRIS on-line Papers in Economics and Informatics*, Vol. 12, No. 1, pp. 61-72. ISSN 1804-1930. DOI 10.7160/aol.2020.120106.

[23] Nagy, J. and Jámbor, Z. (2019) "Competitiveness in Dairy Trade – the Case of EU and the Visegrad Group Countries", *AGRIS on-line Papers in Economics and Informatics*, Vol. 11, No. 4, pp. 61-74. ISSN 1804-1930. DOI 10.7160/aol.2019.110406.

[24] Nowak, A. and Kaminska, A. (2016) "Agricultural competitiveness: The Case of the European Union Countries", *Agriculture Economics – Czech*, Vol. 62, No. 11, pp. 507-516. E-ISSN 1805-9295, ISSN 0139-570X. DOI 10.17221/133/2015-AGRICECON.

[25] Pokrivčák, J., Záhorský, T. and Svetlanská, T. (2015) "Evaluation of Technical Efficiency and Productivity of Agriculture in EU Member States", *Proceedings of 10th International Conference on Applied Business Research*. [Online]. Available: http://www.icabr.com/fullpapers/icabr2015.pdf. [Accessed: Sept. 14, 2020].

[26] Qineti, A., Rajcaniova, M. and Matejkova, E. (2009) "The competitiveness and comparative advantage of the Slovak and the EU agri-food trade with Russia and Ukraine", *Agriculture Economics – Czech*, Vol. 55, No. 8, pp. 375-383. E-ISSN 1805-9295, ISSN 0139-570X. DOI 10.17221/42/2009-AGRICECON.

[27] Rozkošová, A. and Megyesiová, S. (2017) "Labour Productivity of EU Member States in Selected Setions of Economy", *Proceedings of 11th International Days of Statistics and Economics*, Prague. [Online]. Available: https://msed.vse.cz/msed_2017/article/255-Rozkosova-Anna-paper.pdf. [Accessed: Oct. 15, 2020].

[28] Rungsutiyawiboon,S. and Lissitsa, A. (2007) "Agricultural Productivity Growth in the European Union and Transition Countries", In Fuller, F. H. and Beghin, J. C. "*European Agriculture – Enlargement, Structural Change, CAP Reform, and Trade Liberalization",* Nova Science Publishers. ISBN 978-1-60021-966-5.

[29] Ružeková, V., Kittová, Z. and Steinhauser, D. (2020) "Export Performance as a Measurement of Competitiveness", *Journal of Competitiveness*, Vol. 12, No. 1, pp. 145-160. E-ISSN 1804-1728. ISSN 1804-171X. DOI 10.7441/joc.2020.01.09.

[30] Serrao, A. (2003a) "A comparison of Agricultural Productivity among European Countries", *New Medit*, Vol. 2, No. 1, pp. 14-20. ISSN 15945685.

[31] Serrao, A. (2003b) "Agricultural Productivity Analysis of European Union and Eastern Regions", *AAEA – 2003 Annual Meeting*, Montreal, Canada. 20 p. DOI 10.22004/ag.econ.22167.

[32] Sunge, R. and Ngepah, N. (2020) "The Impact of Agricultural Trade Liberalization on Agricultural Total Factor Productivity Growth in Africa", *International Economic Journal*, Vol. 34, No. 4, pp. 571-598. E- ISSN 1743-517X, ISSN 1016-8737. DOI 10.1080/10168737.2020.1836671.

[33] Tiffin, A. (2014) "*European Productivity, Innovation and Competitiveness: the Case of Italy*", IMF Working Papers. ISBN 9781484379868. DOI 10.5089/9781484379868.001.

[34] Tong Soo, K. (2013) "The Gains from External Scale Economies and Comparative Advantage" *Economics Bulletin*, AccessEcon, Vol. 34, No. 1, pp. 84-88.

[35] Torgerson, W. S. (1952) "Multidimensional scaling: I. Theory and method", *Psychometrika*, Vol. 17, No. 4, pp. 401-419. E-ISSN 1860-0980, ISSN 0033-3123. DOI 10.1007/BF02288916.

[36] Ward, J. H. J. (1963) "Hierarchical Grouping to Optimize an Objective Function", *Journal of American Statistical Association*, Vol. 58, No. 301, pp. 236-244. DOI 10.1080/01621459.1963.10500845.

[37] Wulong, G. and Beiling, Y. (2017) "Productivity Growth and International Competitiveness", *Review of Income and Wealth*, No. 63, Vol. 1. ISSN 1475-4991. DOI 10.1111/roiw.12254.

Agris on-line Papers in Economics and Informatics

Volume XII Number 4, 2020

Land Pricing Model: Price Re-evaluation Due to the Erosion and Climate Change Effects

Josef Slaboch, Lukáš Čechura

Department of Economics, Faculty of Economics and Management, Czech University of Life Sciences Prague, Czech Republic

Abstract

The aim of this study is to derive and apply the hedonic approach for determining and updating official land prices with respect to e.g. the impact of climate change that has occurred in the conditions of the Czech Republic in recent years. Pricing using the hedonic method is based on capturing individual factors separately. The evaluated soil ecological unit code consists of a 5-digit numerical code, which expresses the affiliation to the climate region (0-9, see table 1), the main soil unit (0-78), the slope of the land and the orientation to the point of the compass (0-9) and also the depth of the soil profile and skeletality (0-9). The derived hedonic pricing model is estimated using heteroscedasticity corrected estimator. The fitted model shows considerably high explanatory power and together with high parameter significance for majority of dummy variables (soil characteristics) as well as with theoretical and logical consistency represent a tool for new official land price settings in the process of land reevaluation due to the erosion and climate change effects.

Keywords

Soil, hedonic price, land policy, evaluated soil ecological unit.

Slaboch, J. and Čechura, L. (2020) "Land Pricing Model: Price Re-evaluation Due to the Erosion and Climate Change Effects", *AGRIS on-line Papers in Economics and Informatics*, Vol. 12, No. 4, pp. 111-121. ISSN 1804-1930. DOI 10.7160/aol.2020.120408.

Introduction

Agricultural land is generally classified together with labour and capital as the fundamental factor of production for the functioning of a market economy. In this respect, soil is a very specific production factor, as its properties do not allow reproduction, relocation and at the same time the soil has a limited extent. For the above reasons, it is therefore necessary to protect this production resource in a certain way for future generations, as the landscape is the most important element of multifunctional agriculture (Cahill, 2001).

Setting the right policy for the use of individual production resources is a key factor for the economic and social development of individual countries. Land tenure and protection also play a key role in the sustainable development of the countryside and rural areas (Schwarcz et al., 2013).

An important issue is the correct setting of the land price, as it also serves to determine the tax liability of real estate. Many studies around the world have dealt with pricing and the influence of factors on price.

The value of a land usually reflects the quality of the soil, qualitative characteristics, which, however, relate only to the agricultural use. The results of mathematical models explaining the differences in land prices among countries suggest that almost two-thirds of these are attributed to non-agricultural uses, which can significantly distort the production function (Peterson, 1986). These findings are confirmed by recent studies, where the results lead to a similar conclusion, revealing that the agricultural land prices are only partially explained by agricultural yields (production function), and other non-production factors also enter the land price (Borchers, Ifft, and Kuethe, 2014). For these reasons, Garcia and Grande (2003) suggest the use of statistical techniques and methods based on multidimensional analysis to value agricultural land, which refines and simplifies the identification of variables involved in the valuation of agricultural land.

If we focus on the production factors influencing the price of agricultural land, then e.g., Nickerson et al. (2012) showed a positive correlation between soil quality and soil price. Based on the division of land into three categories (high, medium and low quality), it can be quantified that medium quality soils were 5% cheaper than high quality soils. Cotteleer et al. (2007) add that agricultural land prices are also affected by land fragmentation.

According to Kocur-Bera (2016), the key variables influencing the price of agricultural land are the location of land in relation to rural settlements, soil quality, soil fragmentation, forest cover and also the location of farms in less-favoured areas (LFA) for agricultural production.

In Poland, the results show the key importance of areas of rural functional types (agritourism areas), which received on average 43% higher prices than in conventional agricultural areas. In addition, the results show that payments within the LFA area and agri-environmental payments decapitalize land value (Czyzewski, Przekota and Poczta-Wajda, 2017). Gelan and Schwarz (2008) point out that in Scotland, single payments have significantly greater negative effects (with respect to land price) on farms in LFA areas than on farms located outside these areas, and local circumstances need to be taken into account when adjusting the common agricultural policy (CAP).

Another factor influencing the price of land is the access to transport infrastructure. In their study, Sheng, Jackson and Lawson (2018) or Eagle et al. (2015) state that a 1% reduction in transport costs between farmers and ports leads to a 0.33% increase in land prices and there is no significant difference between modes of transport. (road, rail, etc.). Similar conclusions are reached by Woch et al. (2011), when the distance of agricultural land from rural settlements reduces farm incomes and thus indirectly reduces the price of agricultural land. Cavailhès and Thomas (2013) in a survey of Belgian municipalities (589 municipalities) conclude that the price of agricultural land is falling by 2.5% for each km of distance from the municipality.

At a time of reducing total emissions, the results of a land use study are also interesting. The results show that agricultural land built up by solar panels increases the prices of the surrounding agricultural land in the range of 3.4% -37% (depending on the distance from the solar panels) compared to the average market price of agricultural land in the area. In developing countries, this fact also has negative consequences, as potential land tenants or buyers cannot afford to rent/buy agricultural land at a higher price (Lai Mei-Chun et al., 2019).

In Sweden, Nilsson and Johansson (2013) analysed the determinants of agricultural land prices with a focus on area-specific factors. The results show that for areas with low land prices, CAP (single payments for farmers) subsidies have the highest impact. The impact of subsidies under the CAP is also confirmed by Kocur-Bera (2016), where individual subsidies increase the prices of agricultural property, and therefore agricultural land. This conclusion is also reached by Latruffe and Mouel (2009), who state that the aid increases the price of agricultural land (and rent), which favours landowners over agricultural producers.

Agricultural land has been subject to increasing intensification in recent years and it is necessary to prevent deterioration of soil quality. The basic theoretical premise is that landowners have a higher motivation to maintain the quality of land than its tenants. Daedlow, Lemke and Helming (2018) did not demonstrate this premise, and concluded that no clear relationship could be distinguished between the variables used in the model.

Climate change (for example temperature, precipitation), which has occurred in recent decades, can have very serious consequences for agricultural production. For some areas, on the contrary, climate change is an opportunity. Schmitz et al. (2014) point to the fact that 7 out of 10 scenarios in their study assume an increase in fertile soil of 10-25% compared to 2005 (only 1 scenario assumes a decrease). In all models, areas are expanding in South America and sub-Saharan Africa.

In recent years, drought and dryness have become key issues on a global scale (especially in Europe), mainly due to environmental and socio-economic impacts. Drought reduces the biological and economic productivity of ecosystems. Salvati et al. (2012) point to the fact that during the period 2004-2007, more than 50% of the surveyed areas were classified as dry, compared to 0% for the period 2000-2003.

Bozzola et al. (2017) point to the fact that farm incomes are very sensitive to seasonal changes in temperature and precipitation, whereas farms with an irrigation system have the main advantage in this regard. Similar results are obtained by Hossain et al. (2019), when farmers' incomes are influenced mainly by the temperature and the ability of agricultural holdings to use irrigation equipment. The results show that the implementation of global models on climate change can have an impact on income growth in this area, namely by 25-84 USD per hectare.

Chatzopoulos and Lippert (2015) examine the impact of climate change on land prices and assess the impact on individual farm types. Some findings are interesting from the results - with higher temperatures, permanent crops predominate, while in areas with higher precipitation, fodder crops dominates. Land rental prices (lease fees) show concave reactions to the growth of annual precipitation and at the same time increase linearly with increasing temperature. Due to the expected rise in temperatures, climate change is beneficial for most farmers, with the exception of feed producing farmers.

Belyaeva and Bokusheva (2018) analyse the impact of climate change on cereal production in the Russian Federation. For some areas (northern and Siberian), the results show a positive impact of climate change on production. However, due to the high negative impact on the most productive areas in the south of the country, the overall impact of climate change is assessed as negative.

The significant rise in agricultural land prices has led to discussions about the need for more intensive interventions in agricultural land markets. However, changed or new interventions in agricultural land markets should be based on previous analyses of the factors causing the price differences (Lehn and Bahrs, 2018). Ferguson, Furtan and Carlberg (2006) state that one of the ubiquitous forms of agricultural regulation is the restriction of ownership of agricultural land. An example of a restriction on agricultural land ownership is, for example, the Saskatchewan Farmers' Safety Act in Canada. The regulation reduced the prices of Saskatchewan farmland by an average of $ 4 to $ 34 / acre in 1974.

As already mentioned, the price of agricultural land also serves to determine the tax liability. In this case, it is necessary to distinguish two basic types of prices - market and official. Official prices are important in determining the production potential of specific soils in different areas with different natural, ecological or environmental conditions. As stated by Bradáčová (2007), the official price of land enters into property and fiscal relations, but also in soil protection.

In some countries, the official price of land, which is lower than the market price, is used to reduce the tax liability. In the USA, for example, the difference between the tax liability of agricultural entities using the official price and the market price is estimated at USD 60 million. The official price of land is given by the cadastral value - the official price is used, for example, in Austria, Belgium, Denmark, Latvia, Switzerland, the Czech Republic or in the USA. A more detailed comparison of individual EU countries shows that the obligations of individual agricultural entities are very different in terms of calculation methodology, applied rates or what is actually the subject of taxation (OECD, 2019).

Exemptions from land or agricultural property tax is in practise, for example, in the following countries: Australia, Canada (some provinces), Finland, Italy, Japan, Slovenia (excluding farm buildings), Sweden, and the United Kingdom. In some countries, on the other hand, only buildings used for agricultural production are exempt from tax or the tax is significantly reduced - for example, France, Hungary, Latvia, Lithuania, Norway, Poland, etc. Alternatively, reduced tax rates may be applied to farmers for calculation of tax liability from land or agricultural buildings - these are, for example, countries: Belgium, the Czech Republic, Denmark, Estonia, France, etc. A more detailed comparison of the Czech Republic and selected neighbouring countries shows a significant difference in the calculation of tax liability for agricultural entities. In the Czech Republic, the tax base for agricultural land is determined by the product of the acreage of land and the price determined by Decree No. 288/2018 Coll., which sets the list of cadastral areas with assigned average land prices. If we multiply the tax base by a percentage rate of 0.25% - 0.75% (depending on the type of agricultural land), we get the resulting tax liability.

In Poland, real estate tax is calculated as part of the so-called agricultural tax. The rate depends on the average purchase prices of rye, which will be announced by the President of the Statistical Office within 20 days after the end of the 3rd quarter. The municipal council may, by decree, reduce the purchase price of rye for the area where it is locally competent to determine the agricultural tax. According to the law, the tax base is a land size that exceeds 1 hectare and is either real or recalculated depending on whether the land is included in the agricultural holding or not. The recalculated area depends on the so-called conversion coefficients, which are determined depending on the type of land, the classification of the land into a class and the location of the land in one of the 4 districts. The amount of rye depends on whether it is land of agricultural or not. For agricultural land, the rate of rye is 2.5 q (quintal) and for non-agricultural land, the rate is higher and is 5 q of rye.

In Germany, property tax is governed by the Grundsteuergesetz (GrStG). According to § 2 of the GrStG, agricultural and forestry land, buildings and units in Germany are subject to real estate tax. German tax legislation distinguishes between real estate tax A (agrarian), which is levied on agricultural land, and real estate tax B (construction), which applies to built-up or buildable land and buildings. As in the Czech Republic, the law does not distinguish between property owned by a natural or legal person. The tax base is determined on the basis of property valuation and is determined by the local authority. The valuation of assets is determined in accordance with different valuation regulations for different types of assets. A tax rate of 6% applies to agricultural and forestry enterprises.

Materials and methods

The aim of this study is to apply the hedonic approach for determining and updating official land prices with respect to e.g. the impact of climate change that has occurred in the conditions of the Czech Republic in recent years. Another need for the robust tool for setting the official land price is the introduction of new land codes and more detailed land price stratification.

The supporting data are based on a comprehensive soil survey (1960-1970), which is a combination of soil survey and agronomic survey to determine nutrient levels, soil reactions, etc. The comprehensive soil survey is then based on the evaluation of soils, where the goal is the economic valuation of individual factors of the area (climate, relief, soil unit). The valuation is based on different production and cost assumptions of individual types of agricultural land, which are expressed by evaluated soil ecological unit (ESEU) - which is the basic unit of agricultural land valuation. The basis of the ESEU valuation is the creation of valuation type structures (VTS), which express the shares of the appropriate representation of the most important crops on arable land, on the basis of which the individual ESEU codes are valued using the scoring method. The economic valuation itself is calculated on the basis of the gross annual rent effect (GARE), which represents the difference between revenues and costs in the parameterized crop production in a given VTS. There are currently 2318 defined ESEU codes for the Czech Republic. From the above, the obsolescence of the data on which the entire calculation is based is evident - despite minor updates. The determination of climate regions does not currently meet the criteria - usually it refers to a higher average temperature than that assigned by the methodology to the climate region and a lower total annual precipitation than defined in the climate regions.

Pricing using the hedonic method consists in separating the individual factors entering into the final pricing. The ESEU code consists of a 5-digit numerical code, which expresses the affiliation to the climate region (0-9, see table 1), the main soil unit (0-78), the slope of the land and the orientation to the point of the compass (0-9) and also the depth of the soil profile and skeletality (0-9). In the second phase, a hedonic econometric model was specified using these variables. A total of 2,172 ESEU codes were used in accordance with applicable legislation (Decree No. 441/2013 Coll.), the rest of the codes have not yet been valued within the conditions of the Czech Republic. The data for the calculation and application of the hedonic method are based on Act no. 441/2013 Coll. (annex no. 1) – there are the prices of individual ESEU valid for the given period. The hedonic model will be used for valuation of other ESEU, that have not yet been valued (there is no valuation for 146 code).

Model specification:

$$Y_i = f(K, SDR, SDRH, D) \qquad (1)$$

where Y_i is a price of ESEU (CZK/m^2), K stands for a vector of dummy variables on climate region, SDR is a vector of dummies on the combine characteristics of the land slope and exposition, SDRH represents a vector of dummies for the combine effect of the depth of soil profile and skeletality and D is a vector of dummy variables representing the main soil unit. The detail variable specification is provided in Table 1.

We apply a heteroscedastic corrected linear regression model to estimate parameters of model (1).

Dummy variable	Specification
K1	Climate region 1 – warm, dry
K2	Climate region 2 – warm, slightly dry
K3	Climate region 3 – warm, slightly moist
K4	Climate region 4 – slightly warm, dry
K5	Climate region 5 – slightly warm, slightly humid
K6	Climate region 6 – slightly warm, warm, very humid
K7	Climate region 7 – slightly warm, humid
K8	Climate region 8 – slightly cold, humid
K9	Climate region 9 – cold, moist
SDR1	Land slope: 3-7 degrees, exposition: without exposition
SDR2	Land slope: 3-7 degrees, exposition: south
SDR3	Land slope: 3-7 degrees, exposition: north
SDR4	Land slope: 7-12 degrees, exposition: south
SDR5	Land slope: 7-12 degrees, exposition: north
SDR6	Land slope: 12-17 degrees, exposition: south
SDR7	Land slope: 12-17 degrees, exposition: north
SDR8	Land slope: 17-25 degrees, exposition: south
SDR9	Land slope: 17-25 degrees, exposition: north
SDRH1	Depth of soil profile: 30/60 cm and more, skeletality: skeletaless to weakly skeletal
SDRH2	Depth of soil profile: 60 cm and more, skeletality: weakly skeletal
SDRH3	Depth of soil profile: 60 cm and more, skeletality: moderately skeletal
SDRH4	Depth of soil profile: 30 cm and more, skeletality: moderately skeletal
SDRH5	Depth of soil profile: less than 30 cm, skeletality: weakly skeletal
SDRH6	Depth of soil profile: less than 30 cm, skeletality: moderately skeletal
SDRH7	Depth of soil profile: 30 cm and more, skeletality: weakly skeletal*
SDRH8	Depth of soil profile: 30 - 60 cm, skeletality: strongly skeletal
SDRH9	Depth of soil profile: 30 - 60 cm, skeletality: moderately skeletal
D_2 till D_76	Dummies for each main soil unit

Note: * applies to soil units with land slope above 12 degrees (soil unit 40, 41)
Source: Author's own processing

Table 1: Model variable specification

Results and discussion

Table 2 provides parameter estimates of hedonic pricing model (1). The majority of fitted parameters are highly significant, even at 1 % significance level. The only exceptions are 9 out of 75 parameters on the main soil unit.

We employed the heteroscedasticity corrected estimator since the original estimate of linear regression model contained the heteroscedastic error structure. The R^2, as a measure of goodness fit, with other model statistical characteristic (Table 3) indicate good statistical properties. In particularly, R^2 = 0.945, shows that 94.5 % of the variability of official land prices is explained by employed dummy variables. The high explanatory power of the hedonic model with high parameter significance for majority of dummies (soil characteristics) are important factors determining the robustness of the official land price model as a tool for new price settings. However, another important model characteristic that must be met by the estimate is the logical consistence of fitted parameters.

The parameters of dummy variables for climate regions show the following patterns. The parameters represent the change in the official land price (CZK/m^2) with respect to the based region, in this case K0 – very warm and dry region. That is, the parameter of K1 – warm and dry region – indicate that the price for this region is lower by 0.62 CZK/m^2 as compared to the K0 region.

Variable	Coefficient	Std.error	p-value		Variable	Coefficient	Std.error	p-value	
const	14.999	0.374	0.000	***	D_25	-5.139	0.391	0.000	***
K1	-0.620	0.081	0.000	***	D_26	-6.143	0.385	0.000	***
K2	0.343	0.084	0.000	***	D_27	-6.344	0.391	0.000	***
K3	1.398	0.089	0.000	***	D_28	-5.131	0.390	0.000	***
K4	-0.885	0.082	0.000	***	D_29	-5.992	0.390	0.000	***
K5	-0.202	0.078	0.000	***	D_30	-5.672	0.394	0.000	***
K6	-0.300	0.076	0.000	***	D_31	-6.407	0.389	0.000	***
K7	-1.327	0.087	0.000	***	D_32	-6.904	0.393	0.000	***
K8	-1.751	0.126	0.000	***	D_33	-5.430	0.399	0.000	***
K9	-1.305	0.165	0.000	***	D_34	-6.612	0.470	0.000	***
SDR1	-1.055	0.042	0.000	***	D_35	-6.544	0.485	0.000	***
SDR2	-1.360	0.104	0.000	***	D_36	-7.734	0.554	0.000	***
SDR3	-1.331	0.111	0.000	***	D_37	-0.104	0.126	0.407	
SDR4	-2.061	0.044	0.000	***	D_39	0.241	0.409	0.556	
SDR5	-1.996	0.044	0.000	***	D_40	0.044	0.265	0.869	
SDR6	-1.085	0.379	0.004	***	D_41	0.064	0.262	0.806	
SDR7	-1.084	0.404	0.007	***	D_42	-3.921	0.486	0.000	***
SDR8	-1.088	0.450	0.016	**	D_43	-4.878	0.415	0.000	***
SDR9	-1.085	0.496	0.029	**	D_44	-5.486	0.496	0.000	***
SDRH1	-1.090	0.121	0.000	***	D_45	-5.023	0.421	0.000	***
SDRH2	-1.329	0.074	0.000	***	D_46	-6.672	0.406	0.000	***
SDRH3	-1.871	0.077	0.000	***	D_47	-7.513	0.370	0.000	***
SDRH4	-3.425	0.127	0.000	***	D_48	-7.033	0.392	0.000	***
SDRH5	-11.226	0.383	0.000	***	D_49	-7.914	0.390	0.000	***
SDRH6	-11.516	0.386	0.000	***	D_50	-6.836	0.389	0.000	***
SDRH7	-12.365	0.600	0.000	***	D_51	-7.971	0.383	0.000	***
SDRH8	-12.378	0.599	0.000	***	D_52	-7.501	0.389	0.000	***
SDRH9	-12.398	0.568	0.000	***	D_53	-7.491	0.393	0.000	***
D_2	0.174	0.690	0.800		D_54	-8.337	0.386	0.000	***
D_3	1.545	0.940	0.100		D_55	-7.139	0.486	0.000	***
D_4	-6.712	0.418	0.000	***	D_56	-3.549	0.820	0.000	***
D_5	-4.982	0.506	0.000	***	D_57	-5.388	0.961	0.000	***
D_6	-2.942	0.428	0.000	***	D_58	-6.179	0.759	0.000	***
D_7	-3.685	0.555	0.000	***	D_59	-7.418	0.536	0.000	***
D_8	-4.104	0.397	0.000	***	D_60	-1.209	1.053	0.251	
D_9	-0.056	0.750	0.941		D_61	-2.923	1.076	0.007	***
D_10	-0.203	0.465	0.662		D_62	-4.988	0.548	0.000	***
D_11	-1.906	0.426	0.000	***	D_63	-9.568	0.470	0.000	***
D_12	-2.597	0.425	0.000	***	D_64	-7.729	0.405	0.000	***
D_13	-3.939	0.388	0.000	***	D_65	-9.895	0.431	0.000	***
D_14	-2.719	0.396	0.000	***	D_66	-11.918	0.499	0.000	***
D_15	-3.848	0.403	0.000	***	D_67	-12.139	0.450	0.000	***
D_16	-6.051	0.413	0.000	***	D_68	-10.612	0.424	0.000	***
D_17	-7.826	0.382	0.000	***	D_69	-11.943	0.430	0.000	***
D_18	-5.712	0.396	0.000	***	D_70	-10.077	0.422	0.000	***
D_19	-4.970	0.400	0.000	***	D_71	-10.699	0.418	0.000	***
D_20	-6.287	0.386	0.000	***	D_72	-11.887	0.446	0.000	***
D_21	-8.713	0.372	0.000	***	D_73	-10.349	0.400	0.000	***
D_22	-7.679	0.372	0.000	***	D_74	-10.344	0.400	0.000	***
D_23	-7.560	0.381	0.000	***	D_75	-9.847	0.399	0.000	***
D_24	-5.435	0.394	0.000	***	D_76	-9.842	0.400	0.000	***

Note: ***, **, * indicate the level of significance 1%, 5 % or 10%, respectively
Source: Author's estimate

Table 2: Parameter estimate of hedonic pricing model.

Then, the marginal effects of regions K2 – warm and modestly dry - and K3 – warm and modestly moist - are positive, i.e. the marginal prices for these regions are higher by 0.343 and 1.398 CZK/m^2, respectively. The marginal effects for regions K4 till K9 are negative with increasing values from K5 till K8. This estimated marginal prices and the difference among the prices are consistent with our expectations.

Sum squared residuals	6614.4	S.E. of regression	1.788
R-squared	0.945	Adjusted R-squared	0.942
F(102, 2069)	345.2	P-value(F)	0.000

Source: Author's own processing

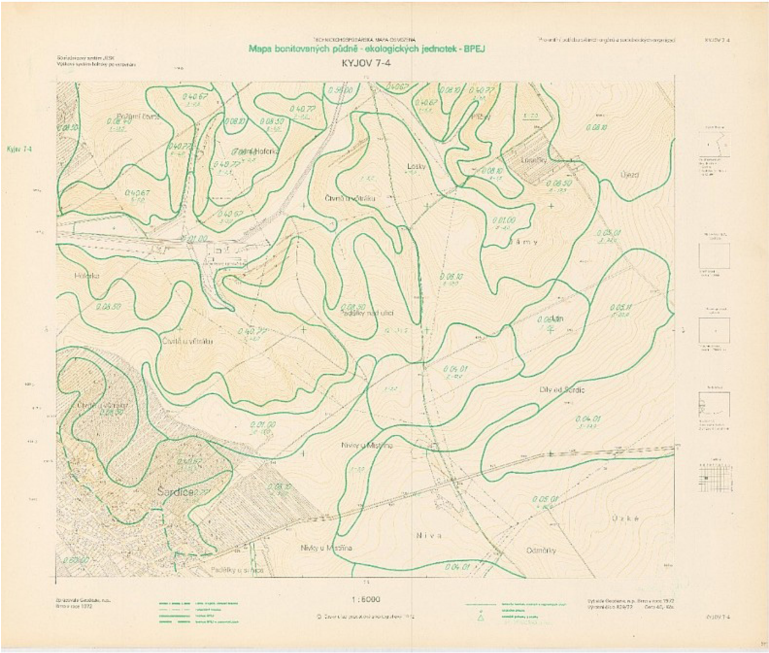
Table 3: Statistical characteristics of fitted model.

The combine characteristics of the land slope and exposition show negative marginal prices with respect to the SDR0 – flatland without exposition. SDR1 – modest slope (3 – 7 degrees) without exposition – has the price by 1.055 CZK/m2 lower as compared to SDR0. Analogically, SDR2 – slope 3-7 degrees, south exposition - has lower price by 1.360 CZK/m^2, SDR3 – slope 3-7 degrees, north exposition - by 1.331 CZK/m^2, SDR4 – slope 7-12 degrees, south exposition by 2.061 CZK/m^2, SDR5 – slope 7-12 degrees, north exposition - by 1.996 CZK/m^2, SDR6 – slope 12-17 degrees, south exposition - by 1.085 CZK/m^2, SDR7 – slope 12-17 degrees, north exposition - by 1.084 CZK/m^2, SDR8 – slope 17-25 degrees, south exposition - by 1.088 CZK/m^2 and finally SDR9 slope 17-25 degrees, north exposition - by 1.085 CZK/m^2. That is, we can observe similar prices for the soils with the same slope and slopes higher than 12 degrees (this is the case of SDR6 till SDR9 with the slope 12 till 25 degrees).

The combine characteristics of the depth of soil profile and skeletality show similar patterns. That is, the estimates indicate negative effects of the soil types as compared to the SDRH0 – depth of soil profile more than 60 cm and skeletaless. In particular, SDRH1 – with depth of soil profile 30 cm and more and skeletaless or weakly skeletal – has the price lower by 1.09 CZK/m2, SDRH2 – with depth of soil more than 60 cm and weakly skeletal - by 1.329 CZK/m^2, SDRH3 – with depth of soil more than 60 cm and moderately skeletal - by 1.871 CZK/m^2, SDRH4 – with depth of soil more than 30 cm and moderately skeletal - by 3.425 CZK/m^2, SDRH5 – with depth of soil less than 30 cm and weakly skeletal - by 11.226 CZK/m^2, SDRH6 – with depth of soil less than 30 cm and moderately skeletal - by 11.516 CZK/m^2, SDRH7 – with depth of soil more than 30 cm and weakly skeletal - by 12.365 CZK/m^2 (applies for soil units with land slope above 12 degree), SDRH8 – with depth of soil more than 30 cm till 60 cm and strongly skeletal - by 12.378 CZK/m^2, and SDRH9 – with depth of soil more than 30 cm till 60 cm and moderately skeletal - by 12.398 CZK/m^2. That is, the soil with SDRH5 till SDRH9 belongs to the group of less quality soils and have considerable lower price. Similar results are achieved by a study that evaluates the impact of certain factors on the agricultural land market (one of the factors is agronomic factors). From the results we can confirm the basic assumption that with the quality of the soil goes up the price (O´Donoghue et al., 2015).

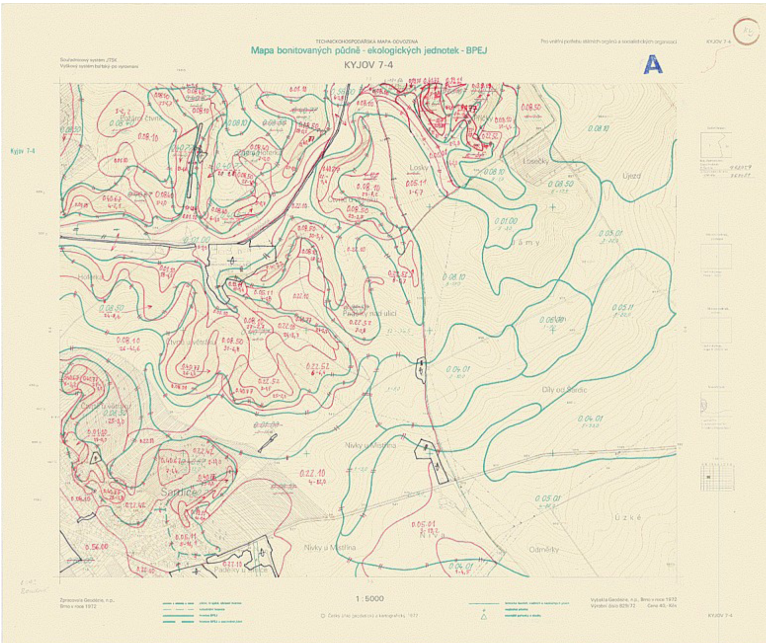
Finally, the parameters of dummy variables for each main soil unit indicate different prices. However, the values are consistent with the assumption about the positive relation between the price and soil quality. The decomposition of the main soil unit value is the subject of the next research activities.

Figure 1 demonstrates ESEU characterization and structure of the selected area, in this case the part of cadastral community in Šardice before actualization in 1973 serves as an example. Figure 2 presents the same area but with ESEU characterization after the actualization in 2000. The figures indicate the different soil characteristics when mapping the area more in detail. This helps to better evaluate the different soil quality with subsequent economic consequences. For this reason and for updating the soil quality in dynamic perspective the robust, solid and unbiased method for price setting is needed. This study is the first and considerably promising attempt to provide such a tool.



Source: VÚMOP

Figure 1: ESEU in cadastral community Šardice – before actualisation in 1973.



Source: VÚMOP

Figure 2: ESEU in cadastral community Šardice – after actualisation in 2000.

Conclusion

The overall good statistical, econometric properties with high explanatory power and especially the economic and logical consistency of fitted parameters suggest that the fitted hedonic pricing model might be a good candidate for the calculation of new official land prices or recalculation of the current prices due to the changes in soil code specification.

Currently, according to the valid legislation in the Czech Republic, there are 2172 ESEU codes. At the same time, in the coming years it is planned to expand the main soil units with other types and to adjust the classification of individual

ESEU into climate regions. In the context of these changes, the presented model is easy to apply due to the very good overall statistics. The advantage of the hedonic approach in determining the price of individual ESEU codes is primarily the speed with which the resulting model can respond to changes in individual influencing parameters. The resulting model serves as a tool for possible recalculation of ESEU prices in the case of a change in the first input parameter, which is defined as a climate region (the climate region is characterized by average temperatures, precipitation, etc.). The use of the resulting econometric model consists in the possible application by the state administration, which is in charge of this issue. At the same time, the valuation problem of individual ESEUs is marginally reflected in the calculation of the official price of land, from which the tax liability (real estate tax) for individual business entities managing agricultural land is subsequently calculated. Due to the nature of these circumstances, it is necessary to verify and sensitively assess the possible impacts on tax collection and the impact on individual private entities.

Acknowledgements

This paper was created within the framework of the project NAZV QK1920280 - Innovation of bonitation system of agricultural lands (ESEU) for the needs of state administration.

Corresponding authors
Ing. Josef Slaboch, Ph.D.
Department of Economics, Faculty of Economics and Management
Czech University of Life Sciences Prague, Kamýcká 129, Prague - Suchdol, 165 00, Czech Republic
Phone: +420 224 382 843, E-mail: jslaboch@pef.czu.cz

References

[1] Belyaeva, M. and Bokusheva, R. (2018) "Will climate change benefit or hurt Russian grain production? A statistical evidence from a panel approach", *Climatic Change,* Vol. 149, No. 2, pp. 205-217. ISSN 1573-1480. DOI 10.1007/s10584-018-2221-3.

[2] Borchers, A., Ifft, J. and Kuethe, T. (2014) "Linking the Price of Agricultural Land to Use Values and Amenities", *American Journal of Agricultural Economics*, Vol. 96, pp. 1307-1320. ISSN 1467-8276. DOI 10.1093/ajae/aau041.

[3] Bozzola, M., Massetti, E., Mendelsohn, R. and Capitanio, F. (2017) "A Ricardian analysis of the impact of climate change on Italian agriculture", *European Review of Agricultural Economics*, Vol. 45, No. 1, pp. 57-79. ISSN 0165-1587. DOI 10.1093/erae/jbx023.

[4] Bradáčová, K. (2007) "Official land price in Slovak Republic", *Agriculture Economics – Czech*, Vol. 53, No. 4, pp. 184-188. ISSN 0139-570X. DOI 10.17221/865-AGRICECON.

[5] Cahill, C. (2001) "Multifunctionality: towards an analytical framework", *Tijdschrift voor Sociaalwetenschappelijk onderzoek in de Landbouw*, Vol. 16, No. 2, pp. 59-71. ISSN 0924-481X.

[6] Cavailhès, J. and Thomas, I. (2013) "Are Agricultural and Developable Land Prices Governed by the Same Spatial Rules? The Case of Belgium", *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, Vol. 61, pp. 439-463. ISSN 1744-7976. DOI 10.1111/j.1744-7976.2012.01265.x.

[7] Cotteleer, G., Stobbe, T. and van Kooten G.C. (2007) "*Farmland conservation in the Netherlands and British Columbia, Canada: a comparative analysis using GIS-based Hedonic Pricing Models*". Resource Economics and Policy Analysis (REPA) Research Group. [Online]. Available: https://web.uvic.ca/~repa/publications/REPA%20working%20papers/WorkingPaper2007-06.pdf [Accessed: 23 Jul. 2020].

[8] Czyżewski, B., Przekota, G. and Poczta-Wajda, A. (2017) "The incidence of agricultural policy on the land market in Poland: Two-dimensional and multilevel analysis", *Land Use Policy*, Vol. 63, pp. 174-185. ISSN 0264-8377. DOI 10.1016/j.landusepol.2017.01.016.

[9] Daedlow, K., Lemke, N. and Helming, K. (2018) "Arable Land Tenancy and Soil Quality in Germany: Contesting Theory with Empirics", *Sustainability*, Vol. 10, No. 8. ISSN 2071-1050. DOI 10.3390/su10082880.

[10] Eagle, A. J., Eagle, D. E., Stobbe, T. E. and van Kooten, G. C. (2015) "Farmland Protection and Agricultural Land Values at the Urban-Rural Fringe: British Columbia's Agricultural Land Reserve", *American Journal of Agricultural Economics*, Vol. 97, pp. 282-298. ISSN 1467-8276. DOI 10.1093/ajae/aau098.

[11] Ferguson, S., Furtan, H. and Carlberg, J. (2006) "The political economy of farmland ownership regulations and land prices", *Agricultural Economics*, Vol. 35, pp. 59-65. ISSN 1574-0862. DOI 10.1111/j.1574-0862.2006.00139.x.

[12] Garcia, T. and Grande, I. (2003) "A model for the valuation of farmland in Spain", *Journal of Property Investment & Finance*, Vol. 21, No. 2, pp. 136-153. ISSN 1463-578X. DOI 10.1108/14635780310469111.

[13] Gelan, A. and Schwarz, G. (2008) "The effect of single farm payments on less favoured areas agriculture in Scotland: a CGE analysis", *Agricultural and Food Science,* Vol. 17, No. 1, pp. 3-17. ISSN 1795-1895. DOI 10.2137/145960608784182317.

[14] Hossain, M., Qian, L., Arshad, M., Shahid, S., Fahad, S. and Akhter, J. (2019) "Climate change and crop farming in Bangladesh: an analysis of economic impacts", *International Journal of Climate Change Strategies and Management*, Vol. 11, No. 3, pp. 424-440. ISSN 1756-8692. DOI 10.1108/IJCCSM-04-2018-0030.

[15] Chatzopoulos, T. and Lippert, C. (2015) "Adaptation and Climate Change Impacts: A Structural Ricardian Analysis of Farm Types in Germany", *Journal of Agricultural Economics*, Vol. 66, pp. 537-554. ISSN 1477-9552. DOI 10.1111/1477-9552.12098.

[16] Kocur-Bera, K. (2016) "Determinants of agricultural land price in Poland – a case study covering a part of the Euroregion Baltic", *Cahiers Agricultures*, Vol. 25. No. 2. ISSN 1166-7699. DOI 10.1051/cagri/2016013.

[17] Lai Mei-Chun, Wu Pei-Ing, Liou Je-Liang, Chen Yi and Chen Hanhui (2019) "The impact of promoting renewable energy in Taiwan - How much hail is added to snow in farmland prices?", *Journal of Cleaner Production,* Vol. 241. ISSN 0959-6526. DOI 10.1016/j.jclepro.2019.118519.

[18] Latruffe, L. and Le Mouël, C. (2009) "Capitalization of government support in agricultural land price: What do we know?", *Journal of Economic Surveys*, Vol. 23, pp. 659-691. ISSN 1467-6419. DOI 10.1111/j.1467-6419.2009.00575.x.

[19] Lehn, F. and Bahrs, E. (2018) "Analysis of factors influencing standard farmland values with regard to stronger interventions in the German farmland market", *Land Use Policy*, Vol. 73, pp. 138-146. ISSN 0264-8377. DOI 10.1016/j.landusepol.2018.01.020.

[20] Nickerson, C., Morehart, M., Kuethe, T., Beckman, J., Ifft, J. and Williams, R. (2012) "Trends in U.S. Farmland Values and Ownership", *AgEcon*, United States Department of Agriculture (USDA), Economic Information Bulletin, Vol. 92, pp. 1-48. DOI 10.22004/ag.econ.291935.

[21] Nilsson, P. and Johansson, S. (2013) "Location determinants of agricultural land prices", *Jahrbuch für Regionalwissenschaft*, Vol. 33, pp. 1-21. ISSN 1613-9836. DOI 10.1007/s10037-012-0071-4.

[22] O´Donoghue, C., Lopez, J., O´Neill, S., Ryan, M. (2015) "A hedonic price model of self-assessed agricultural land values", Conference paper of *EAAE seminar "The spatial dimension in analyzing the linkages between agricultural, rural development and the environment*. [Online]. Available: file:///C:/Users/jslaboch/AppData/Local/Temp/A%20Hedonic%20Price%20Model%20of%20Self-Assessed%20Agricultural%20Land%20Values.pdf [Accessed: 2 Dec. 2020].

[23] OECD (2019) "*Taxation in agriculture*". [Online]. Available: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=TAD/CA/APM/WP(2018)30/FINAL&docLanguage=En [Accessed: 23 Jan. 2020].

[24] Peterson, W. (1986) "Land Quality and Prices", *American Journal of Agricultural Economics*, Vol. 68, pp. 812-819. ISSN 1467-8276. DOI 10.2307/1242127.

[25] Salvati, L., Perini, L., Sabbi, A. and Bajocco, S. (2012) "Climate Aridity and Land Use Changes: A Regional-Scale Analysis", *Geographical Research*, Vol. 50, pp. 193-203. ISSN 1745-5871. DOI 10.1111/j.1745-5871.2011.00723.x.

[26] Sheng, Y., Jackson, T. and Lawson, K. (2018) "Evaluating the benefits from transport infrastructure in agriculture: a hedonic analysis of farmland prices", *Austtralian Journal of Agriculture and Resource Economics*, Vol. 62, pp. 237-255. ISSN 1467-8489. DOI 10.1111/1467-8489.12243.

[27] Schmitz, C., van Meijl, H., Kyle, P., Nelson, G.C., Fujimori, S., Gurgel, A., Havlik, P., Heyhoe, E., d'Croz, D. M., Popp, A., Sands, R., Tabeau, S., van der Mensbrugghe, D., von Lampe, M., Wise, M., Blanc, E., Hasegawa, T., Kavallari, A. and Valin , H. (2014) "Land-use change trajectories up to 2050: insights from a global agro-economic model comparison", *Agricultural Economics*, Vol. 45, pp. 69-84. ISSN 1574-0862. DOI 10.1111/agec.12090.

[28] Schwarcz, P., Bandlerová, A. and Schwarczová, L. (2013) "Selected issues of the agricultural land market in the Slovak Republic", *Journal of Central European Agriculture*, Vol. 14, No. 3, pp. 1102-1115. ISSN 1332-9049. DOI 10.5513/JCEA01/14.3.1314.

[29] VÚMOP (2020) Interníal materials.

[30] Woch, F., Wierzbicki, K., Eymont, A., Dziadkowicz-Ilkowska, A., Syp, A., Kopiński, J., ... (2011) "Economic viability and economic land consolidation in Poland", Puławy: Institute of Soil Science and Plant Cultivation, in Polish.

Agris on-line Papers in Economics and Informatics

Volume XII Number 4, 2020

Application of Markowitz Portfolio Theory to Producing the World Major Field Crops

Natalia Vasylieva

Department of Information Systems and Technologies, Dnipro State Agrarian and Economic University, Ukraine

Abstract

Development of growing cereals and oilseeds is a pressing issue for providing global food security and renewable energy. The study deals with applying methods of portfolio theory to mitigate natural and marketing uncertainties emerged from unstable yields and volatile prices for wheat, maize, barley, sunflower, soybeans, and rapeseed. The research outcome based on the utilization of Markowitz mean-variance indicators made possible to evaluate portfolio performances of the world top cereals and oilseeds producers. The study findings at a country level combined econometric forecasting of the crop revenues and modeling optimal portfolios of cereals and oilseeds subject to acceptable trade-offs between risks and expected revenues. The fulfilled calculations with Ukrainian focus clarified farmland allocations under cereal and oilseed crops to underpin biodiversity and keep firm positions in the world markets.

Keywords

Cereals and oilseeds; risk; production portfolio; revenue forecast; Markowitz models.

Vasylieva, N. (2020) "Application of Markowitz Portfolio Theory to Producing the World Major Field Crops", *AGRIS on-line Papers in Economics and Informatics*, Vol. 12, No. 4, pp. 123-131. ISSN 1804-1930. DOI 10.7160/aol.2020.120409.

Introduction

The world major field crops are cereals and oilseeds. They are utilized for food, feed, and bio-fuel production. As of 2019, the prime cereal crops like wheat, rice, maize, and barley occupied 214.8, 167.1, 193.7, and 48 million hectares of farmland around the world and were cultivated in 123, 117, 168, and 104 countries respectively. Similarly, the prime oilseed crops such as sunflower, soybeans, and rapeseed covered 26.7, 124.9, and 37.6 million hectares of the world arable land in 73, 99, and 65 countries (FAOStat, 2020).

At the same time, agriculture is a risky business. It implies that farmers have to meet the challenges of output uncertainty caused by the natural factors and price volatility affected by fluctuations of the market supply and demand. The most promising tool to facilitate this issue is a portfolio approach. In case of the agricultural sector, portfolio production can measure a risk, predict revenue, and allocate a restricted resource of farmland on different assets like cash crops. Developing such optimal win-win strategies will comply with the environmental, social, and economic goals of maintaining agricultural biodiversity, providing the global food security in view of the accelerated population growth, and gaining competitive advantages in the domestic and international agricultural markets. Given the pressing aspects of portfolio production of the world major field crops, it is worth a separate thorough scientific study.

Creating risk-efficient operating plans and strategies is a key point in contemporary agriculture. Farmers around the globe are exposed to production, marketing, financial, legal, and personnel risks. Science proposes a wide variety of means and tools to handle such uncertainties. But the most fruitful risk-management decisions are those adjusted to specific agricultural products (Broll et al., 2013).

The European cropping pattern includes the above listed world major cereals and oilseeds except for rice. Given the Ukrainian research focus, this investigation considered wheat, maize, barley, sunflower, soybeans, and rapeseed. Advanced economics of their production, processing, marketing, and utilization was revealed in numerous comprehensive studies, in particular by Carver (2009), Danforth (2011), Elfson (2011), Martinez-Force et al. (2015), Johnson et al. (2015), and Gunstone (2004).

Villanueva et al. (2017) clarified that the witnessed biodiversity within the groups of cereals and oilseeds have a positive impact on crops resilience and adaptability to the ongoing climate change. Besides, Brussaard et al. (2010), Frison et al. (2011) indicated importance of biodiversity in providing food security via preventing declines in yields of the core staple crops. In case of Ukraine, biodiversity contributed to robust food security by cereals and oilseeds, unlike other industries of the national agriculture (Vasylieva, 2018; Vasylieva, 2019).

Mathematical models developed by Brummer et al. (2016), Haile et al. (2016), Santeramo and Lamonaca (2019) offered accurate and convincing evidence that risks in supplies dependent on yield fluctuations strongly affect price volatility in cereals and oilseeds markets. In other words, it is crucial to allocate farmland under a mix of crops to match both natural and market conditions (Skrypnyk et al., 2018; Ramankutty et al., 2018).

In this regard, the most appropreate mathematical apparatus is the portfolio concept intended to assist in selecting weights of assets to mitigate possible losses (Kolm et al., 2014). Such approach demonstrates reliable results in modelling diversified crop production. In particular, based on the assumption that "typically, risk in the farm income arises from risk in revenue", Mumey et al. (1992, p. 71) developed models of reducing revenue risk experienced by the wheat, barley, and canola producers at the local level. Barkley et al. (2010) examined a portfolio of wheat varieties directed at increasing yields, shrinking risk and enhancing profitability. Radulescu et al. (2014) elaborated optimal portfolios under minimum environmental and financial risks or maximum expected returns from growing wheat, corn, barley, sunflower, and rapeseed. Recent findings on this topic included models on mean-variance planning for crop farms (Toth et al., 2016).

However, the world production portfolios of the most demanded cereals and oilseeds need an updated comparison which entails options of the optimal crop portfolios applicable at a country level. In light of the agricultural resources and potentials, the latter was conducted for Ukraine. With this research objective, the study was divided into three tasks:

- to implement Markowitz mean-variance indicators to track progress in performing production portfolios of the world major cereals and oilseeds;
- to forecast fluctuations of revenues from wheat, maize, barley, sunflower, soybeans, and rapeseed caused by unstable yields and price volatility in Ukraine; and
- to utilize Markowitz models for cereals and oilseeds to calculate their production portfolios subject to an acceptable trade-off between risk and expected revenue.

Materials and methods

To cope with the task 1, this research employed the mean-variance fundamentals of Markowitz portfolio theory (Markowitz, 2010). In more detail, N was a quantity of crops in an evaluated portfolio;

W_i designated a portfolio share of each crop subject to $W_i \geq 0$, $i = 1,\dots,N$, $\sum_{i=1}^{N} W_i = 1$;

T defined an analyzed time frame; and

X_t^i referred to an annual revenue from crop *i* at a point in time *t*, *i* = 1...*N*, *t* = 1,..., *T*.

It made possible to compare portfolio performance among the world top crop producers via expected portfolio revenue (EPR) and portfolio risk (PR) built on the logarithmic indices of crop revenue relative to values in the preceding year:

$$Z_t^i = \mathrm{Ln}(X_t^i / X_{t-1}^i),\ t = 2,\dots,T,$$

$$Z^i = \sum_{t=2}^{T} Z_t^i / (T-1),\ i = 1,\dots,N.$$

In this fashion, the named portfolio indicators were calculated as follows:

$$EPR = \sum_{i=1}^{N} W_i \cdot Z^i, \quad (1)$$

$$PR = \sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N}(W_i \cdot W_j \cdot \sum_{t=2}^{T}((Z_t^i - Z^i)\cdot(Z_t^j - Z^j))/(T-1))}. \quad (2)$$

In addition, an indicator of the present portfolio revenue (PPR) illuminated a current output in absolute terms like

$$PPR = \sum_{i=1}^{N} W_i \cdot X_T^i. \quad (3)$$

The panel data intended for the task 1 involved time-series of revenues from the world major portfolio crops by top countries.

A methodical background to the task 2 was econometrics (Greene, 2007; Studenmund, 2016). It was focused on capturing trend and cycle components of dynamics in crop revenue. For such reason, the offered regression incorporated linear and sinusoidal components such as

$$X(t) = A_0 + A_1 \cdot t + A_2 \cdot \sin(A_3 + A_4 \cdot t). \quad (4)$$

Here *X(t)* denoted a calculated crop revenue (in $ per hectare) associated with a time variable *t*. The numerical regression coefficients A_0 - A_4 were found through the least squares method applied to the time-series data concerning the analyzed major field crops, i.e.

$$\sum_{t=1}^{T}(X_t - (A_0 + A_1 \cdot t + A_2 \cdot \sin(A_3 + A_4 \cdot t)))^2 \to \min. \quad (5)$$

As before, *T* denoted a time frame, and X_t was a real crop revenue referred to the observation *t*. The regression coefficients allowed comparing dynamics in different crops. Namely, the coefficient A_1 described an annual change in a revenue trend. Meanwhile, the coefficient A_2 revealed the revenue volatility via a cycle amplitude.

R-squared and F-test verified the regression applicability to forecasting crop revenue over the future periods. In case of inadequate accuracy of calculations, the utilized least squares method should be replaced by its advanced modification.

To arrange a trade-off between the expected portfolio revenue and portfolio risk, the research task 3 dealt with both maximum and minimum Markowitz models (Prigent, 2007). In compliance with the previous identifications, W_1, $i = 1, \ldots, N$ were the model variables associated with the sought shares of crops in production portfolios. The given lower bounds to the share of each crop enabled to promote biodiversity assuming that

$$S_i > 0, \ i = 1, \ldots, N, \text{ and } \sum_{i=1}^{N} S_i < 1.$$

The models input data were obtained from the tasks 1 and 2. In particular, the values of EPR_{min} and PR_{max} introduced the acceptable levels of minimal expected portfolio revenue (1) and maximal portfolio risk (2).

Thus, the maximum Markowitz model translated into

$$EPR \to \max \quad (6)$$

subject to $$W_i \geq S_i, \ i = 1, \ldots, N, \ \sum_{i=1}^{N} W_i = 1, \quad (7)$$

$$PR \leq PR_{max}. \quad (8)$$

The minimum Markowitz model looked like

$$PR \to \min \quad (9)$$

subject to $$W_i \geq S_i, \ i = 1, \ldots, N, \ \sum_{i=1}^{N} W_i = 1, \quad (10)$$

$$EPR \geq EPR_{min}. \quad (11)$$

Applications of the formulated Markowitz models under different scenarios can ground the farmers' strategies on diminishing portfolio risks in times of recession or increasing expected portfolio revenues in times of economic expansion. As such, the calculated shifts in production portfolios will provide the national agricultural authorities with the objective forecasts about country ranks among the world top rivals in cereals and oilseeds markets.

Results and discussion

The described research methodology entailed the relevant empirical outcomes presented in this section.

Global crop portfolios

Hereafter there was intended for wheat, maize, and barley in the group of cereals or for sunflower, soybeans, and rapeseed in the group of oilseeds. To synchronize outcomes of the task 1 with further results referring to Ukraine, the study time frame covered the period of 1996 to 2018 started from launching Ukrainian national currency. It stipulated $T = 23$.

Consistent with Mumey et al. (1992), the presented research examined a portfolio risk in expected revenue. To define the latter, a comprehensive panel data to the task 1 contained yields and farmgate prices for wheat, maize, barley, sunflower, soybeans, and rapeseed by country (FAOStat, 2020). The conducted world comparison considered countries which practice production portfolios and occupy over 1% of the total harvested area under at least 2 crops amongst the analyzed cereals and oilseeds.

As such, Table 1 encompassed 20 countries which were engaged in cereals production portfolios and accumulated 83.1%, 64.3%, and 78.8% of the global farmland under wheat, maize, and barley respectively. Overall, it means a strong commitment of cereals growers to implementing portfolios. Besides, Indonesia, Nigeria, and the Philippines also practiced crop portfolios cultivating 9.6%, 2%, and 2.9% of the total farmland under rice as well as 2.9%, 2.5%, and 1.3% of the world arable land under maize. Similarly, Pakistan occurred to comprise 1.7% of rice and 4.1% of wheat total harvested areas. But given the rice components, the described portfolios were beyond the study focus.

A thorough insight on the indicators in Table 1 revealed different features and priorities of the listed countries. Indeed, China, India, Ethiopia, and Iran are the 1st, 2nd, 12th, and 18th countries by population. In light of providing food security to meet high steady domestic demands, they demonstrated relatively low risks and high absolute revenues. It is also worth mentioning that India and China are the top growers of rice with 26.6% and 18.1% of its total harvested area. In compliance with Rude and An (2015), the largest cereals exporters such as France, Germany, and the USA had moderate portfolio risks paired with average expected portfolio revenues offset by high revenues in absolute terms. The Argentinean, Brazilian, Canadian and Ukrainian portfolio performance appeared to be quiet similar and promising. To some extent, Poland, Romania, Spain and Turkey illustrated opposite priorities in accepting risks justified by higher revenues. In spite of large scale cereals production, Australia, Kazakhstan, and Russia were merged by risky and inefficient results. The worst portfolio performance was found in Algeria and Morocco which should revise their agricultural strategies and improve portfolio indicators of EPR (1), PR (2), and PPR (3).

Table 2 reported data about 10 countries which were involved in oilseeds production portfolios and accumulated 74.6%, 63.1%, and 75.4% of the world harvested areas under sunflower, soybeans, and rapeseed respectively.

Table 2 displayed that the oilseeds portfolio of China dominated over that of India by most indicators. The top world exporters such as France and the USA had similarly balanced oilseeds portfolios. The strongest risk acceptance emerged in Romania that presented an opposite operational priority compared to Canada. The Ukrainian oilseeds portfolio showed better relative indicators of EPR (1) and PR (2) than those in Argentina but, unfortunately, did not gain by PPR (3). As before, despite large occupied farmlands, Kazakhstan and Russia had risky and inefficient oilseeds portfolios in absolute terms.

Country	Share (%) in the World Farmland under			Expected Portfolio Revenue (%)	Portfolio Risk (%)	Present Portfolio Revenue ($/ha)
	Wheat	Maize	Barley			
Algeria	1.0	0.0	2.7	1.8	29.4	581.3
Argentina	2.7	3.7	2.5	2.1	22.1	801.6
Australia	5.1	0.0	8.6	0.2	31.5	414.5
Brazil	1.0	8.3	0.2	3.4	22.1	765.5
Canada	4.6	0.7	5.0	2.9	23.6	693.9
China	11.3	21.7	0.8	1.8	21.6	1596.4
Ethiopia	0.8	1.2	2.0	5.5	20.6	711.0
France	2.4	0.7	3.7	0.8	21.2	1369.0
Germany	1.4	0.2	3.4	0.6	24.8	1292.6
India	13.8	4.7	1.4	5.8	14.7	906.2
Iran	3.1	0.1	3.1	3.6	13.4	704.2
Kazakhstan	5.3	0.1	5.2	2.9	41.5	151.2
Morocco	1.3	0.1	3.3	2.0	52.0	637.6
Poland	1.1	0.3	2.0	0.5	26.4	794.9
Romania	1.0	1.3	0.9	5.1	42.5	1105.2
Russia	12.3	1.2	16.4	3.0	34.2	366.7
Spain	1.0	0.2	5.4	1.6	37.3	887.1
the USA	7.5	17.1	1.7	1.4	16.6	1209.6
Turkey	3.4	0.3	5.4	1.1	18.6	594.9
Ukraine	3.1	2.4	5.2	5.4	26.6	785.6

Source: own calculation based on FAOStat (2020)

Table 1: Evaluation of cereals portfolios by top countries.

Country	Share (%) in the World Farmland under			Expected Portfolio Revenue (%)	Portfolio Risk (%)	Present Portfolio Revenue ($/ha)
	Sunflower	Soybeans	Rapeseed			
Argentina	6.3	13.1	0.0	2.3	21.3	829.1
Canada	0.1	2.0	24.3	3.1	20.5	883.1
China	3.3	6.4	17.4	3.0	15.0	1244.6
France	2.1	0.1	4.3	1.7	18.5	1119.2
India	1.0	9.1	17.8	4.4	19.5	722.5
Kazakhstan	3.2	0.1	1.0	8.1	24.8	301.4
Romania	3.8	0.1	1.7	8.7	32.8	980.9
Russia	29.8	2.2	4.0	6.1	21.2	470.0
the USA	1.9	28.5	2.1	2.9	13.9	1167.6
Ukraine	23.1	1.4	2.8	7.2	16.3	814.8

Source: own calculation based on FAOStat (2020)

Table 2: Evaluation of oilseeds portfolios by top countries.

Revenue forecast

As of 2019, Ukraine ranked the 6th, 3rd, 4th, 1st, 6th, and 3rd among the world top exporters of wheat, maize, barley, sunflower oil, soybeans, and rapeseed. These achievements are essential incentives to delve deeper into improving production portfolios of cereals and oilseeds at the national scope.

The data of six time-series to the task 2 were the revenues (in $ per hectare) from the explored crops delivered by the State Statistics Service of Ukraine (2020) for 1996 to 2019. Hereafter it meant that . The coefficients of the regressions (4) for wheat, maize, barley, sunflower, soybeans, and rapeseed were found by means of the least squares method (5) and assembled in Table 3.

These findings combined and expanded econometric models developed separately for crop prices by Vasylieva (2013) and for yields by Arunachalam and Balakrishnan (2012). As it can be seen from Table 3, all considered cereals and oilseeds had upward revenue trends. The coefficient   identified the least annual increase in the revenue for barley and the highest growth in the revenue for rapeseed. Oilseeds appeared to be more uniform by the range of the forecasted revenues for 2020. According to the coefficients A_2, cereals had more stable revenues except for maize. In all cases, the values of R-squared and F-significance permitted applying the forecasted crop revenues to the next stage of this study concerning the optimal production portfolios for cereals and oilseeds in Ukraine.

Portfolio optimization

As mentioned before, the calculations to the task 3 stipulated $N = 3$ and $T = 23$. Given the biodiversity encouragement, parameters S_1, S_2 and S_3 accounted for 0.05 that implied allocating of at least 5% of farmland under each crop in an appropriate portfolio. The conducted calculations were linked with four the most beneficial scenarios. Namely,

- scenario #1 supposed portfolio production with the risk below its current level in Ukraine, to wit, $PR_{max} = 0.266$ for cereals (see Table 1) and $PR_{max} = 0.163$ for oilseeds (see Table 2) in the model (6)-(8);
- scenario #2 addressed portfolio production without restrictions to the risk level, i.e. running the reduced model (6), (7) for both groups of crops in question;
- scenario #3 suggested portfolio production with the expected revenue over its current level in Ukraine, to wit, $EPR_{min} = 0.054$ for cereals (see Table 1) and $EPR_{min} = 0.072$ for oilseeds (see Table 2) in the model (9)-(11);
- scenario #4 defined portfolio production without restrictions to the expected revenue level, i.e. running the reduced model (9), (10) for both groups of crops in question.

The computed figures were aggregated in Tables 4 and 5. Their analysis resulted in the following inferences.

Crop	A_0	A_1	A_2	A_3	A_4	R-squared	F-significance	Revenue Forecast for 2020 ($/ha)
Cereals								
Wheat	95.41	23.09	60.10	-0.55	1.16	0.87	0.00	658
Maize	52.56	42.57	135.97	1.45	0.39	0.88	0.00	987
Barley	39.77	19.00	51.22	-0.10	1.16	0.86	0.00	487
Oilseeds								
Sunflower	-9.29	34.51	113.88	1.63	-0.37	0.95	0.00	744
Soybeans	66.32	34.16	88.21	2.21	-0.40	0.91	0.00	833
Rapeseed	-7.89	46.55	120.90	1.33	-0.37	0.94	0.00	1035

Source: own calculation based on State Statistics Service of Ukraine (2020)

Table 3: Output to regressions on crop revenue.

Scenario	Share (%) in the World Farmland under			Expected Portfolio Revenue (%)	Portfolio Risk (%)
	Wheat	Maize	Barley		
#1	35	59	6	5.5	26.6
#2	5	90	5	5.9	30.5
#3	30	63	7	5.4	26.5
#4	25	59	16	4.5	24.7

Source: own calculation based on State Statistics Service of Ukraine (2020)

Table 4: Optimal portfolios for cereals.

Scenario	Share (%) in the World Farmland under			Expected Portfolio Revenue (%)	Portfolio Risk (%)
	Sunflower	Soybeans	Rapeseed		
#1	38	5	57	7.4	16.3
#2	5	5	90	7.8	22.4
#3	47	5	48	7.2	15.4
#4	48	21	31	6.8	14.5

Source: own calculation based on State Statistics Service of Ukraine (2020)

Table 5: Optimal portfolios for oilseeds.

Firstly, both cereals and oilseeds portfolios derived from the scenarios #2 gravitated to more specialized productions. According to Czyzewski and Smedzik-Ambrozy (2015), similar farmers' strategies may be explained by shifting to the crops with larger revenues, reduced technological costs, and diminished marketing expenditures through the deteriorated biodiversity.

Secondly, both cereals and oilseeds portfolios derived from the scenarios #4 fostered more diversified productions driven by the farmers' strategic priority to mitigate natural and marketing risks (Lin, 2011).

Thirdly, under scenarios #1 and #3 Ukraine retained its positions by EPR and PR ranked 3rd and 13th for cereals (see Tables 1 and 4) as well as ranked 3rd twice for oilseeds (see Tables 2 and 5). Scenario #2 applied to cereals elevated Ukrainian EPR up to 1st rank but decreased its PR by 1 position. Scenario #4 acted the opposite way. It dropped Ukrainian EPR to 4th rank but lifted its PR up to 10th rank. Under scenarios #2 and #4 Ukrainian ranks for oilseeds were constant by EPR but shifted to 8th and 2nd by PR.

Finally, according to PPR Ukrainian portfolio revenues were relatively low in absolute terms ranked 10th out of 20 countries by cereals and 7th out of 10 countries by oilseeds. Sure enough, Ukrainian farmers are price-takers in the global scale and can not influence a marketing risk (Velychko et al., 2019). Nevertheless, it is an unfavorable but not a hopeless situation.

Ukrainian farmers are able to increase their yields of cereals and oilseeds to reach the levels of the most advanced producers via improved technologies and implemented innovations (Vasylieva and Pugach, 2017). As such, the national agricultural autorities ought to encourage this activity and incentivize biodiversity through the political measures of the permissible financial protection of the domestic cereals and oilseeds growers.

Conclusion

In sum, the largest portfolio producers cultivate 341 million hectares or 74.7% of the world farmland under wheat, maize, and barley and 127 million hectares or 67.1% of the total harvested areas under sunflower, soybeans, and rapeseed. These crops are a cornerstone in providing food, feed and bio-fuel resources. Markowitz indicators unfolded that the top cereals producers were more numerous and less uniform compared to the top oilseeds growers. The latter ones practiced less risky activity with higher expected portfolio revenues. The research insight also revealed that Argentina, Canada, China, France, India, Kazakhstan, Romania, Russia, the USA, and Ukraine run large scale portfolios by both cereals and oilseeds. As of 2019, their combined cultivated area amounted to 388.9 million hectares or 60.2% of the world farmland under these crops.

A combination of econometric and optimization models appeared to be a relevant mathematical technique to enhance portfolio performance by the world major field crops at a country level. It made possible to alleviate uncertainty in yields and prices triggered by natural and marketing risks intrinsic to agribusiness. The proposed utilization of Markowitz models delivered options on reasonable farmland allocation and biodiversity promotion which are global imperatives of contemporary agriculture. With regard to Ukrainian circumstances, the conducted calculations detected opportunities to decrease portfolio risks and raise expected portfolio revenues for the considered crops.

Taking into account the annual revenue fluctuations in cereals and oilseeds markets, the recommended portfolios need regular updates and amendments that can define a promising avenue for further scientific elaborations.

Corresponding authors
Natalia Vasylieva, Doctor in Economics, Professor
Department of Information Systems and Technologies, Dnipro State Agrarian and Economic University, 25 Serhiya Yefremova Street, Dnipro, 49600, Ukraine
Phone: +38-098-033-4900, E-mail: VasylievaN@i.ua
ORCID: 0000-0003-4100-0659

References

[1] Arunachalam, R. and Balakrishnan, V. (2012) "Statistical Modeling for Wheat (Triticum Aestivum) Crop Production", *International Journal of Statistics and Applications*, Vol. 2, No. 4, pp. 40-46. ISSN 2168-5193, E-ISSN 2168-5215. DOI 10.5923/j.statistics.20120204.03.

[2] Barkley, A. P., Peterson, H. and Shroyer, J. (2010) "Wheat Variety Selection to Maximize Returns and Minimize Risk: An Application of Portfolio Theory", *Journal of Agricultural and Applied Economics*, Vol. 42, No. 1, pp. 39-55. ISSN 1074-0708, E-ISSN 2056-7405. DOI 10.1017/S107407080000328X.

[3] Broll, U., Welzel, P. and Wong, K. P. (2013) "Price Risk and Risk Management in Agriculture", *Contemporary Economics*, Vol. 7, No. 2, pp. 17-20. ISSN 2084-0845, E-ISSN 2300-8814. DOI 10.5709/ce.1897-9254.79.

[4] Brummer B., Korn O., Schlusler K. and Jaghdani T. J. (2016) "Volatility in oilseeds and vegetable oils markets: Drivers and spillovers", J*ournal of Agricultural Economics*, Vol. 67, No. 3, pp. 685-705. E-ISSN 1477-9552. DOI 10.1111/1477-9552.12141.

[5] Brussaard, L., Caron P., Campbell, B., Lipper, L., Rabbinge, R., Babin, D. and Pulleman, M. (2010) "Reconciling biodiversity conservation and food security: scientific challenges for a new agriculture", *Current Opinion in Environmental Sustainability*, Vol. 2, No. 1-2, pp. 34-42. ISSN 1877-3435. DOI 10.1016/j.cosust.2010.03.007.

[6] Carver, B. F. (2009) "*Wheat: Science and Trade*", Hoboken, New Jersey: Wiley-Blackwell, 616 p. ISBN 978-0813820248.

[7] Czyzewski, A. and Smedzik-Ambrozy, K. (2015) "Specialization and diversification of agricultural production in the light of sustainable development", *Journal of International Studies*, Vol. 8, No 2, pp. 63-73. ISSN 2071-8330, E-ISSN 2306-3483. DOI 10.14254/2071-8330.2015/8-2/6.

[8] Danforth, A. T. (2011) "*Corn Crop Production: Growth, Fertilization and Yield*", Hauppauge, New York: Nova Science Publishers, Inc., 377 p. ISBN 978-1607419556.

[9] Elfson, S. B. (2011) "*Barley: Production, Cultivation and Uses*", Hauppauge, New York: Nova Science Publishers, Inc., 341 p. ISBN 978-1612097459.

[10] FAOStat (2020) "*Data of the Food and Agriculture Organization of the United Nations*", Statistics, Feb. 2020. [Online]. Available: http://www.fao.org/faostat/en/#data [Accessed: 16 Feb. 2020].

[11] Frison, E. A., Cherfas J. and Hodgkin T. (2011) "Agricultural Biodiversity is Essential for a Sustainable Improvement in Food and Nutrition Security", *Sustainability*, Vol. 3, No. 1, pp. 238-253. ISSN 2071-1050. DOI 10.3390/su3010238.

[12] Santeramo F. G. and Lamonaca, E. (2019) "On the drivers of global grain price volatility: an empirical investigation", *Agricultural Economics – Czech*, Vol. 65, No. 1, pp. 31-42. ISSN 0139-570X, E-ISSN 1805-9295. DOI 10.17221/76/2018-AGRICECON.

[13] Greene, W. H. (2007) "*Econometric analysis*", Upper Saddle River, New Jersey: Pearson Prentice Hall, p. 1216. ISBN 978-0135132456.

[14] Gunstone, F. (2004) "*Rapeseed and Canola Oil: Production, Processing, Properties and Uses*", Hoboken, New Jersey: Wiley-Blackwell, p. 240. ISBN 978-1405116251.

[15] Haile, M. G., Kalkuhl, M. and von Braun, J. (2016) "Worldwide acreage and yield response to international price change and volatility: a dynamic panel data analysis for wheat, rice, corn, and soybeans", *American Journal of Agricultural Economics*, Vol. 98, No. 1, pp. 172-190. ISSN 0002-9092, E-ISSN 1467-8276. DOI 10.1093/ajae/aav013.

[16] Johnson, L. A., White, P. J. and Galloway, R. (2015) "*Soybeans: Chemistry, Production, Processing, and Utilization*", New York City, New York: Academic Press and AOCS Press, 850 p. ISBN 978-1893997646.

[17] Kolm, P. N., Tutuncu, R. and Fabozzi, F. J. (2014) "60 Years of portfolio optimization: Practical challenges and current trends", *European Journal of Operational Research*, Vol. 234, No. 2, pp. 356-371. ISSN 0377-2217. DOI 10.1016/j.ejor.2013.10.060.

[18] Lin, B. B. (2011) "Resilience in Agriculture through Crop Diversification: Adaptive Management for Environmental Change", *BioScience*, Vol. 61, No. 3, pp. 183-193. ISSN 0006-3568, E-ISSN 1525-3244. DOI 10.1525/bio.2011.61.3.4.

[19] Markowitz, H. M. (2010) "Portfolio Theory: As I Still See It", *Annual Review of Financial Economics*, Vol. 2, No. 1, pp. 1-23. E-ISSN 1941-1375. DOI 10.1146/annurev-financial-011110-134602.

[20] Martinez-Force, E., Dunford, N. T. and Salas, J. J. (2015) "*Sunflower: Chemistry, Production, Processing, and Utilization*", New York City, New York: Academic Press and AOCS Press, 725 p. ISBN 978-1893997943.

[21] Mumey, G., Burden, B. and Boyda, A. (1992) "Measurement of Farm Risk: Alberta Crop Production", *Canadian Journal of Agricultural Economics*, Vol. 40, No. 1, pp. 71-91. E-ISSN 1744-7976. DOI 10.1111/j.1744-7976.1992.tb03678.x.

[22] Prigent, J.-L. (2007) "*Portfolio Optimization and Performance Analysis*", London: Chapman and Hall/CRC, p. 456. ISBN 978-1584885788. DOI 10.1201/9781420010930.

[23] Radulescu, M., Radulescu, C. Z. and Zbaganu, G. (2014) "A portfolio theory approach to crop planning under environmental constraints", *Annals of Operations Research*, Vol. 219, pp. 243-264. ISSN 0254-5330, E-ISSN 1572-9338. DOI 10.1007/s10479-011-0902-7.

[24] Ramankutty, N., Mehrabi, Z., Waha, K., Jarvis, L., Kremen, C., Herrero, M., and Rieseberg, L. H. (2018) “Trends in Global Agricultural Land Use: Implications for Environmental Health and Food Security“, *Annual Review of Plant Biology*, Vol. 69, pp. 789-815. E-ISSN 1545-2123. DOI 10.1146/annurev-arplant-042817-040256.

[25] Rude, J. and An, H. (2015) “Explaining grain and oilseed price volatility: The role of export restrictions“, *Food Policy*, Vol. 57, No. 11, pp. 83-92. ISSN 0306-9192. DOI 10.1016/j.foodpol.2015.09.002.

[26] Santeramo F. G. and Lamonaca E. (2019) “On the drivers of global grain price volatility: an empirical investigation“, *Agricultural Economics – Czech*, Vol. 65, No. 1, pp. 31-42. ISSN 0139-570X, E-ISSN 1805-9295. DOI 10.17221/76/2018-AGRICECON.

[27] Skrypnyk, A., Tkachuk, V., Andruschenko, V. and Bukin, E. (2018) “Sustainable development facets: farmland and market demand estimation“, *Journal of Security and Sustainability Issues*, Vol. 7, No. 3, pp. 513-525. ISSN 2029-7017, E-ISSN 2029-7025. DOI 10.9770/jssi.2018.7.3(11).

[28] State Statistics Service of Ukraine (2020) “*Agriculture in Ukraine*“. Statistics, Feb. 2020. [Online]. Available: http://www.ukrstat.gov.ua/ [Accessed: 12 Feb. 2020].

[29] Studenmund, A. H. (2016) “*Using Econometrics: A Practical Guide*“, London: Pearson plc., 576 p. ISBN 978-0134182742.

[30] Toth, M., Holubek, I. and Serences, R. (2016) “Applying Markowitz portfolio theory to measure the systematic risk in agriculture”, in *Proceedings of Conference: International Scientific Days 2016. The Agri-Food Value Chain: Challenges for Natural Resources Management and Society*, Slovak University of Agriculture in Nitra, pp. 985-993. DOI 10.15414/isd2016.s12.10.

[31] Vasylieva, N. (2013) “Forecasting of prices in the field of crops growing in Ukraine and regions“, *Economic Annals-XXI*, No. 11-12(2), pp. 26-29. ISSN 1728-6220, E-ISSN 1728-6239.

[32] Vasylieva, N. and Pugach, A. (2017) “Economic assessment of technical maintenance in grain production of Ukrainian agriculture“, *Bulgarian Journal of Agricultural Science*, Vol. 23, No. 2, pp. 198-203. ISSN 1310-0351, E-ISSN 2534-983X.

[33] Vasylieva, N. (2018) “Ukrainian Agricultural Contribution to the World Food Security: Economic Problems and Prospects“, *Montenegrin Journal of Economics*, Vol. 14, No. 4, pp. 215-224. ISSN 1800-5845, E-ISSN 1800-6698. DOI 10.14254/1800-5845/2018.14-4.15.

[34] Vasylieva, N. (2019) “Problems and Prospects of Food Security in Ukraine“, *Bulgarian Journal of Agricultural Science*, Vol. 25, No. 4, pp. 668-676. ISSN 1310-0351, E-ISSN 2534-983X.

[35] Velychko, O., Velychko, L., Butko, O. and Khalatur, S. (2019) “Modelling of strategic managerial decisions in the system of marketing logistics of enterprise“, *Innovative Marketing*, Vol. 15, No. 2, pp. 58-70. ISSN 1814-2427, E-ISSN 1816-6326. DOI 10.21511/im.15(2).2019.05.

[36] Villanueva, A. B., Halewood, M. and Noriega, I. L. (2017) “Agricultural Biodiversity in Climate Change Adaptation Planning“, *European Journal of Sustainable Development*, Vol. 6, No. 2, pp. 1-8. ISSN 2239-5938, E-ISSN 2239-6101. DOI 10.14207/ejsd.2017.v6n2p1.

Agris on-line Papers in Economics and Informatics

Volume XII Number 4, 2020

'Products Mapping' of South Africa's Agri-food trade with the EU28 and Africa

Ivo Zdráhal[1], Nahanga Verter[1], Francois Lategan[2]

[1] Department of Regional and Business Economics, Mendel University in Brno, Czech Republic

[2] Department of Agriculture and Game Ranch Management, Nelson Mandela University, Port Elizabeth, Republic of South Africa

Abstract

This paper uses 'products mapping' tool based on the trade balance index (TBI) and Lafay index to investigate trade performance and competitiveness in food items between South Africa (SA) and the EU28 and Africa. The data for this analysis is obtained from the UNCTAD database. SA's agri-food trade balance climbed from $1.5 billion in 2005 to $3.1 billion in 2017. The results support the conclusion that in bilateral trade, certain products have comparative advantages in relation to African markets despite comparative disadvantages in relation to the EU28 market. Also, there is no or decreasing diversification towards more and new leading products despite the increased intra-regional openness. Leading products (especially fruit and nuts) are the dominant export generating segments in the product's structure of SA's agri-food trade. Also, leading products mostly contribute to the positive balance of SA's agri-food trade. The findings of this study may contribute to business strategies, trade policies, and regional and inter-regional integration.

Keywords

Agri-food, comparative advantage, competitiveness, TBI, Lafay index.

Zdráhal, I., Verter, N. and Lategan, F. (2020) 'Products Mapping' of South Africa's Agri-food trade with the EU28 and Africa, *AGRIS on-line Papers in Economics and Informatics*, Vol. 12, No. 4, pp. 133-149. ISSN 1804-1930. DOI 10.7160/aol.2020.120410.

Introduction

Since 1994, the agricultural sector of South Africa (SA) has undergone substantial policy reforms. The country's economy has diversified, while the share of agricultural GDP declined from 11% to 2.3% between 1960 and 2017 (World Bank, 2019). Nevertheless, the agricultural sector's declining share of GDP did not mean that the sector was declining; it was more an indication that the services sectors grew faster (Vink and Rooyen, 2009). At the same time, the sector has been highly exposed to global markets as farmers do not receive any subsidies, and trade at the borders has been substantially liberalized (Sandrey et al., 2011).

Remarkably, SA's population grew from 41.4 million to 58.6 million between 1995 and 2019. Also, the country has achieved food security status, with 67.3/100 scores (the Global Food Security Index), and ranked 48/113 countries in 2019 (EIU, 2019). Nevertheless, the country still faced some challenges, such as climate shocks in rainfall (FSIN, 2019), market access and modern farm inputs (FAO, 2019; UNCTAD, 2019). Despite the challenges, SA has developed its food sector relative to the sub-region of Sub Saharan Africa (SSA).

The European Union (EU) is a traditional and important market for SA's agri-food exports. However, the EU food market has become saturated and SA faces competition from other southern hemisphere countries with similar seasonal differences in comparison to the northern hemisphere (Adriaen et al., 2004). After the end of apartheid, the EU was the first regional body that SA entered trade negotiations to re-integrate into the global trading system (Larsén, 2007; European Commission, 2016). In 1999, the EU and SA reached the Trade, Development and Co-operation Agreement (TDCA) that was signed in 2000 after four years of rigorous negotiations. The TDCA established a preferential trade arrangement and partially introduced a free trade area agreement (FTA), covering 90% of bilateral trade between SA and the EU. To protect the vulnerable sectors of SA and the EU, certain products were excluded from the FTA, while some of the products partially liberalized. For the EU, these are mostly agricultural products, while for SA, there are industrial products and certain textile and clothing products. In general, the TDCA appeared asymmetrical in favour of SA, but the country felt that there was an unjustified imbalance against its agricultural sector. The country argued that some rules governing agricultural trade were too rigid and should be relaxed (Berends, 2016).

To further strengthen trade relations, establish close and sustain relations based on cooperation and partnership, SA signed the 'EU-Southern African Development Community Economic Partnership Agreement States' (SADC EPA states) together with five other southern African countries. The TDCA ought to be replaced by the EPA once ratified by the member states (European Commission, 2016). To boost market access, the EU recognizes SA as the leading trading partner in the region, and the significance of the agricultural sector in poverty alleviation strategies in the SADC EPA States (Berends, 2016). On food security measures, the Parties recognize that the removal of restrictions to trade between them, as envisaged in the Agreement, might pose major challenges to the SADC EPA States' producers in the food sectors. The Parties, therefore, agreed to consult with each other on these mentioned issues (European Commission, 2016).

In addition to developing Africa's relations with overseas countries and markets, the Pan-Africanism and regional cooperation have been seen as one of the most important instruments to promote economic growth and development in Africa since the period of decolonization (Blížkovský et al., 2018). The SA's is a member of the Southern African Customs Union (SACU). SACU was established already in 1910 making it the world's oldest Customs Union. In 2002, SACU Agreement highlighted, among others, the facilitation of cross-border movement of goods, the promotion of fair competition, and the equitable sharing of customs and excise revenue raised by all Member States within the Union (DTI, 2020).

Also, the regional trade, especially between South Africa and other SADC countries, has grown rapidly since the mid-2000s and has now reached levels that imply considerable macroeconomic significance. Africa, driven principally by SADC, has become the largest destination for diversified manufactured exports from SA (Arndt and Roberts, 2018).

Against this background, the article an attempt to contribute to the development of literature on trade policies and regional trade in the agricultural sector. The main aim of the study is to investigate agri-food trade performance and competitiveness in SA with the EU28, Africa and the world using "products mapping" technique. The study identifies the main agri-food products that positively or negatively contribute to the over-all agri-food trade of SA. The findings of this study may contribute to business strategies, national development and trade policies, and regional integration.

Literature review and some empirical evidence

Assessing the competitiveness of the agri-food industries in the context of global or regional competition, models following Ricardo and his theory suggest that the countries should focus on producing food products with comparative advantage. According to the Heckscher-Ohlin model, the trade specialization pattern is formed based on countries relative endowment in production factors (Nazarczuk et al., 2018). These findings are in line with current researches of productivity factors structure and effective usage (Bilan et al., 2020; Maris, 2019).

Other streams of theoretical literature emphasize the endogeneity of technological change (Krugman, 1987; Lucas, 1988; Brodzicky and Kwiatkowski, 2018; Cieślik, 2018) or economic geography that underlines the importance of agglomeration economies (Krugman, 1991; Fujita et al., 1999; Kostiukevych et al., 2020). Porter (1990) developed the diamond model, suggesting factors (input) conditions, demand conditions, supporting and related industries and corporate strategy, structure and competition as the driving forces of competitiveness of a nation or industry in the global competition. Moreover, it appears that attention should be directed from costs and production efficiencies towards promoting productivity growth over time and innovation (Yang et al., 2019). Removing tariffs on goods traded between countries and reducing nontariff barriers by harmonizing product standards and simplifying government formalities reduces the transaction costs of trade which should lead to an increase in the degree of specialization (Aiginger, 2001). Higher specialization can lead to higher productivity and competitiveness (and vice versa). Generally, the trade theories give dissimilar predictions regarding specialization dynamics of a country.

Some scholars have used comparative advantage approach to investigate the level of agri-food trade performance, comparative advantage and competitiveness (e.g., Fertő, 2008; Bojnec and Fertő, 2015; Smutka et al, 2016; Benesova et al., 2017; Esquivias, 2017; Smutka et al. 2019). Empirical studies on international agri-food trade, comparative advantage and competitiveness suggest and support that changes in trade patterns and performance are due to both demand and supply sides, both at domestic and international markets, both in factor-intensities and productivity differentials. Liberalization, integration, and industrialization are also channels for improvements in productivity, scale, and export expansion and a way to improve comparative advantage.

In recent years, the level of agri-food trade performance, comparative advantage and competitiveness of SA agrarian trade have been investigated. DAFF (2011) measures the trade competitiveness of SA with the EU27 in some agri-food products for the period 2001 and 2009. Using RCA and comparative export performance (CEP) index, the results reveal that SA has been competitive in the EU27 in products, such as fish and crustaceans, fruits and beverages and vegetables. On the other hand, findings show SA with comparative disadvantages in cereals, tobacco, and sugar.

In the same direction, De Pablo Valenciano et al. (2017) investigate trade competitiveness in SA pear (fruit) with the EU28. Their findings show that SA has been highly competitive in pear exports to the EU28. They argue that South Africa's competitive advantage is driven by the trade agreements (TDCA) signed with the EU28.

There are also empirical studies in agri-food trade (Ishchukova and Smutka, 2014; Benesova et al., 2017; Esquivias, 2017; Smutka et. al., 2018; Ortikov et al., 2019, Zdráhal et al., 2019; Verter et al., 2020) that employ the Product mapping technique to analyze both comparative advantage and global competitiveness and its implications for domestic agri-food trade-balance creation. The composition of trade structure developed from Product Mapping and its trajectory give other important insights regarding the country's integration into the global or regional agribusiness.

The Product Mapping' method was designed by Widodo (2009) to analyze the catching-up countries' comparative advantage and its leading exported products. As pointed by Widodo (2009), leading exported products usually have a high comparative advantage. In the same manner, leading exported products spur export and contribute to the domestic trade-balance, are a source of output growth, and foreign exchange earnings.

Ishchukova and Smutka (2014) analysed the agri-food trade dynamic of Russia in relations to EU, Commonwealth of Independent States, Africa, Asia and the Americas. They found that in bilateral trade products of a specific group have comparative advantages in relation to a region or country despite comparative disadvantages in relation to the whole world. Also, Benesova et al. (2017) used product mapping to analysed Russian agri-food trade and concluded that there exists a general trend of strengthening comparative advantages of Russian agricultural exports, because the results of the product mapping method identified a growing share of class A in the total value of Russian agri-food exports and, at the same time, identify a reduction in the proportion of class D.

Esquivias (2017) used product mapping to analyse changes in agricultural trade patterns of East Java, Indonesia versus six main ASEAN exporting countries from 2007 to 2013 and found that gains appear to be larger than the losses; however, there is a little diversification towards new products despite the increased international openness. Also, he concludes, that opportunities within the region have not been exploited, because agri-food trade in the region still concentrates towards extra-ASEAN territories.

Smutka et al. (2018) studied agri-food trade of the Czech Republic from 2001 to 2015. The results derived from product mapping shows that the number of products located in classes B and C has significantly reduced and the whole commodity structure is divided into classes A and D. They point that the over-all product's classification is influenced by the bilateral trade relations and differences in trade regimes.

Ortikov et al. (2019) use the mapping to analyse Uzbek's trade in agricultural products and foodstuffs from 1995 to 2015. They found that agricultural exports of Uzbekistan are competitive especially in relation to the Asian countries and CIS countries. On the other hand, the comparative advantages in relation to other territories are limited.

Zdráhal et al. (2019) use products mapping to analyse agri-food trade between Nigeria and ECOWAS member countries. The findings suggest that Nigeria has performed better in trading with other ECOWAS countries than in trading with the overall world market and the product mapping revealed some of the promising product groups for expansion within the region and potential for Nigeria to diversify its agri-food export structure. In other studies, Verter et al. (2020) indicate that the share of total Nigerian food exports and imports which the EU28 accounted for, declined from 72% and 40% to 37% and 27% between 1995 and 2017, respectively. Also, the food products that Nigeria has comparative disadvantages and negative trade balance in trading with the EU28 rose from 31/46 to 35/46.

Although scholars have used 'products mapping' approach to analyse agri-food competitiveness, to the of our knowledge, no study has used this tool to investigate trade competitiveness in SA. Thus, this study is an attempt to bridge the knowledge gap.

Materials and methods

This paper analyses the dynamics in the food trade of SA with the EU28 and African countries using time series for the period 2005-2017. The data for this analysis is obtained from the UN Conference on Trade and Development (UNCTAD) database. The classification of specific food products used in this paper is adapted from UNCTAD following the UN Standard International Trade Classification (SITC, Revision 3). The values are calculated (current prices, US$) at the three-digit level of the SITC for all the 46 food items (SITC 0 + 1 + 22 + 4) as presented in Table A1 (in appendix).

The Coefficient of concentration and Herfindahl-Hirschman Index (HHI) measures the concentration and competitiveness of countries or regions across the globe (Reis and Farole, 2012). The product concentration index shows how exports and imports of a nation or region concentrate on a few products or otherwise distributed in a more homogeneous manner among a broad range of products. In other words, the index measures the dispersion of export's or import's values across exporter's or importer's products (*n* products). The model is mathematically presented here as follows (Blažková and Chmelíková, 2016):

$$s_l = \frac{x_{ij}}{\sum_1^n x}, \text{ where } l = 3,5,\ 10 \text{ most traded products} \quad (1)$$

The *HHI* is mathematically presented here as follows:

$$HHI = \sum {s_{ij}}^2 \quad (2)$$

Where: s is the share of exports (import) in the total food trade for the product i in the year j between SA and the World, EU28 or Africa. The value of the index ranges from 0 to 1. A value closer to 1 indicates that food trade is concentrated in few goods and/or sectors for trade. Thus, its vulnerability to trade shocks, whereas a thoroughly diversified portfolio will have an index close to 0, suggesting a lesser vulnerability risk. The HHI can be classified as an indication of diversification in the exporter's profile.

To capture the degree of trade specialization of a country, it is also essential to assess the revealed comparative advantages of the relevant sectors included in the total agrarian trade. For this purpose, Lafay index (LFI) is selected (Lafay, 1992). Contrary to the traditional Balassa index that uses only export data to investigate comparative advantage in countries, the LFI uses both export and import data (Benešová et al., 2018). Another advantage of the LFI is its reliability when comparing its values in time series (Sanidas and Shin, 2010). The index is defined for a given country and a product as follows:

$$LFI_j^i = 100\left(\frac{x_j^i - m_j^i}{x_j^i + m_j^i} - \frac{\sum_{j=1}^N \left(x_j^i - m_j^i\right)}{\sum_{j=1}^N \left(x_j^i + m_j^i\right)}\right)\frac{x_j^i + m_j^i}{\sum_{i=j}^N \left(x_j^i + m_j^i\right)} \quad (3)$$

Where: *x* and *m* are the export and import values of individual product group of agrarian trade of SA to/from EU28 countries as well as Africa as a whole. Zero represents a neutral value regarding reporting a comparative advantage. A positive value for the LFI indicates the existence of comparative advantage for a specific sector and a negative value of the LFI indicates the existence of a comparative disadvantage for a sector. This means that a higher index value suggests a higher degree of comparative advantage and specialisation (Zaghini, 2003). The values of LFI were calculated for 46 different products constituting agrarian foreign trade of SA with the different regions.

Empirically, a country might have a comparative advantage for a product, but the country is not a net exporter. Similarly, a country may have a comparative disadvantage but is not a net importer.

The Trade Balance Index (TBI) is employed to analyse whether a nation has achieved advanced levels of specialisation in export (as net-exporter) or import (as net-importer) for a specific group of products. TBI is mathematically formulated as follows:

$$TBI_j^i = \frac{x_j^i - m_j^i}{x_j^i + m_j^i} \quad (4)$$

Where: TBI_{ij} denotes trade balance index of country *i* for product *j*; x_{ij} and m_{ij} represent exports and imports of product products *j* by nation *i*, respectively. The values of the index range from -1 to +1. In extreme cases, the TBI will equal -1 if a nation only imports. On the other hand, the TBI could equal +1 if a nation only exports. Understandably, the index is not defined when a nation neither exports nor imports. A nation is referred to as "net-importer" or consumer of particular product group if the value of TBI is negative. On the contrary, a nation is known as a "net-exporter" of a product if the value of TBI is positive.

Accordingly, the LFI and TBI are combined to create an analytical tool, called 'products mapping'. Similarly, the mapping classifies a product and a country into four categories (Widodo, 2008) as follows: Group A signifies that SA has a comparative advantage and is a net-exporter; Group B signifies that SA has a comparative advantage but is net-importer; Group C signifies that SA has a comparative disadvantage but is a net-exporter; Group D signifies that SA has a comparative disadvantage and is net-importer (Table 1). The technique has been used recently to study agrarian trade of countries in Europe and Asia, and Nigeria in Africa.

Results and discussion

The SA's agri-food trade gained its momentum in the early 2000s when the implementation of the Uruguay Round Agreement on Agriculture was completed. Together with the decrease in unilaterally applied tariffs levels and preferential applied tariffs levels, SA opened its markets as well as got market access and increased its integration in global agribusiness. The total food exports in SA rose from about $4 billion in 2005 to $9.6 billion in 2017 (Table 2).

The country trade balance also climbed from $1.5 billion to $3.1 billion within the same period under review. During the same period, SA recorded a positive balance of trade and TBI in the overall food products (Table 2). However, the value of exports and imports stagnated in the last decade. The following section presents South Africa's trade performance and products mapping with EU28 and with Africa in all 46 food items (No. of sectors) based on the calculated LFI, TBI and other descriptive approaches.

Agri-food trade between SA and the EU28

SA is the EU's largest trading partner in Africa, in total merchandise and food trade (UNCTAD, 2019). The EU's ranking of the global leading agri-food trade partners shows that SA was the number 19 top importing markets (with 1.2% share of extra-EU) for the agri-food exported by the EU28 in 2018. Also, SA was the number 11 top global supplying markets (with 2.5% share of extra-EU) for the agri-food imported by the EU28 in 2018 (European Commission, 2019).

The total value of food exports from SA

LFI > 0	Group B Comparative Advantage No Export-Specialization (net - importer) (LFI > 0) and (TBI < 0)	Group A Comparative Advantage Have Export-Specialization (net - exporter) (LFI > 0) and (TBI > 0)
LFI < 0	Group D Comparative Disadvantage No Export-Specialization (net - importer) (LFI < 0) and (TBI < 0)	Group C Comparative Disadvantage Have Export-Specialization (net - exporter) (LFI < 0) and (TBI > 0)
	TBI < 0	TBI > 0
	Trade Balance Index (TBI)	

Source: Widodo, 2008

Table 1: Product mapping scheme.

Indicator/year	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Ex (Mill. US$)	3,990	3,734	4,214	5,429	5,469	8,231	9,017	8,743	9,293	9,463	8,547	8,426	9,612
Im (Mill. US$)	2,451	2,977	4,128	4,589	4,163	5,493	7,015	7,451	6,769	6,266	5,961	6,234	6,525
Bal (Mill. US$)	1,539	758	85	840	1,303	2,738	2,002	1,292	2,524	3,198	2,586	2,192	3,087
TBI	0.24	0.11	0.01	0.08	0.14	0.20	0.12	0.08	0.16	0.20	0.18	0.15	0.19

Notes: Ex = exports; Im = imports; Bal = trade balance
Source: Own composition based on UNCTAD (2019)

Table 2: Total agri-food trade of South Africa (SITC 0+1+22+4).

to the EU28 slightly rose from $1.8 billion in 2005 to $2.6 billion in 2017. Interestingly, SA recorded a positive balance of trade and TBI in food trade with the EU28 throughout the period under review, although the TBI has decreased (Table 3). It suggests that the country's overall competitiveness in agri-food trade with the EU28 has reduced. The share of SA's food exports (% of total food exports) to the EU28 declined from 44.1% to 26.6% between 2005 and 2017, while food imports (% of total food imports) from the union's markets rose 22.7% to 27.4% within the same period under study. This suggests that SA may have diversified its export markets beyond the EU28 while imports concentration from the union's markets accelerated.

The 'products mapping' in Group A indicates that SA's comparative advantages in bilateral trade with the EU28 fluctuated but increased from 14/46 in 2005 to 17/46 in 2008, then, shrank to its lowest in 2014 with 10/46, before rising to 13/46 food products between 2016 and 2017 in bilateral trade with the EU28. These products accounted for about 80% and 7% of the total food exports and imports, respectively. This indicates that, although the country recorded a positive trade balance with the EU28, it has not marginally diversified in exporting food products. Similar studies (within the same period, using the same methodology) carried out by Verter et al. (2020) in Nigeria reveals that the country with 9/46 comparative advantages in agri-food trade with the EU28. This shows that SA performs slightly better than Nigeria in agri-food trade with the EU28.

On the other hand, the products mapping

Indicator/year		2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Export (Mill. US$)		1,759	1,533	1,985	2,231	2,012	2,225	2,229	2,070	2,360	2,362	2,230	2,302	2,556
Import (Mill. US$)		557	706	909	990	1,073	1,322	1,732	1,825	1,780	1,934	1,725	1,673	1,785
Balance (Mill. US$)		1,203	827	1,076	1,241	939	902	498	245	580	429	505	629	770
TBI		0.52	0.37	0.37	0.39	0.30	0.25	0.13	0.06	0.14	0.10	0.13	0.16	0.18
Share on export (%)*		44.1	41.1	47.1	47.1	36.8	27.0	24.7	23.7	25.4	25.0	26.1	27.3	26.6
Share on import (%)*		22.7	23.7	22.0	21.6	25.8	24.1	24.7	24.5	26.3	30.9	28.9	26.8	27.4
EX	CR3	82.0	77.2	80.2	77.8	80.8	81.3	80.9	83.4	84.5	82.6	83.7	81.7	80.8
	CR5	90.9	87.0	88.9	85.7	88.4	88.8	88.0	89.9	90.3	88.2	89.4	87.9	87.2
	CR10	96.5	96.1	96.4	94.9	95.5	96.0	96.1	96.5	96.5	95.4	96.2	95.8	95.5
	HHI	0.29	0.26	0.29	0.26	0.29	0.31	0.31	0.34	0.35	0.34	0.38	0.37	0.37
IM	CR3	43.1	46.7	51.8	51.2	58.7	49.2	50.2	49.4	50.4	44.9	42.8	46.3	37.8
	CR5	52.2	57.1	62.1	62.4	68.7	63.4	64.3	63.3	65.2	58.8	58.6	60.4	55.5
	CR10	71.6	76.4	79.1	80.8	82.8	82.5	82.7	81.4	83.3	79.9	81.2	80.6	79.0
	HHI	0.10	0.11	0.15	0.14	0.15	0.11	0.11	0.10	0.11	0.09	0.09	0.09	0.08
A	No. of products	14	15	14	17	13	12	13	13	13	10	11	13	13
	Export (Mill. US$)	1,687	1,447	1,918	2,124	1,935	2,088	2,132	1,983	2,266	1,705	1,710	2,164	2,035
	Export share (%)	95.9	94.4	96.7	95.2	96.2	93.9	95.6	95.8	96.0	72.2	76.7	94.0	79.6
	Import (Mill. US$)	175	225	406	406	433	355	432	429	463	65	76	346	119
	Import share (%)	31.4	31.9	44.6	41.0	40.3	26.8	24.9	23.5	26.0	3.4	4.4	20.7	6.7
	Balance (Mill. US$)	1,513	1,222	1,513	1,718	1,502	1,733	1,700	1,554	1,804	1,640	1,634	1,818	1,916
C	No. of products	1	1	1	1	1	2	1	0	2	4	1	1	2
	Export (Mill. US$)	0.23	8	0.51	5	3	20	0.19	0	10	545	411	4	423
	Export share (%)	0.0	0.5	0.0	0.2	0.2	0.9	0.0	0.0	0.4	23.1	18.5	0.2	16.6
	Import (Mill. US$)	0.20	7	0.50	5	2	15	0.19	0	9	364	289	3	297
	Import share (%)	0.0	1.0	0.1	0.5	0.2%	1.1	0.0	0.0	0.5	18.8	16.7	0.2	16.6
	Balance (Mill. US$)	22	1	8	240	626	5	4	0	1	181	123	0.94	126
D	No. of products	31	30	31	28	32	32	32	33	31	31	34	32	31
	Export (Mill. US$)	72	78	66	102	74	117	97	88	84	113	108	134	98
	Export share (%)	4.1	5.1	3.3	4.6	3.7	5.2	4.4	4.2	3.6	4.8	4.9	5.8	3.8
	Import (Mill. US$)	382	474	503	580	638	953	1,300	1,397	1,309	1,505	1,360	1,324	1,370
	Import share (%)	68.6	67.1	55.3	58.6	59.4	72.0	75.0	76.5	73.5	77.8	78.8	79.1	76.7
	Balance (Mill. US$)	-310	-396	-437	-477	-564	-836	-1,202	-1,309	-1,225	-1,392	-1,251	-1,191	-1,272

Note: * share of bilateral food trade between SA and the EU28
Source: Own composition based on UNCTAD (2019)

Table 3: Dynamics of food trade between SA and the EU28.

in Group D reveals that SA recorded comparative disadvantages and adverse trade balance in 31/46 between 2005 and 2017 in trading with the EU28. The share of these product groups also fluctuated during the period under review and recorded about 77% and 4% of total imports and exports respectively in 2017. More detailed information on products in groups A and D are presented in Table A5. It suggests the performance of South Africa’s exports to the EU28 has not improved significantly in many food products. Similar studies by Verter et al. (2020) in Nigeria reveals that the country with 35/46 comparative disadvantages in agri-food trade with the EU28. This shows that SA performs slightly better than Nigeria n agri-food trade with the EU28.

A critical look at the individual product groups (Table A2) in bilateral trade with the EU28 shows that throughout the period under review, SA has comparative advantages in products, such as SITC 001 (fish, aqua. invertebrates, prepared, preserved); SITC 016 (fish, aqua. invertebrates, prepared, preserved); SITC 034 (fish, fresh, chilled or frozen); SITC 036 (crustaceans); SITC 037 (fish, aqua. invertebrates, prepared, preserved); SITC 057 (fruits and nuts), SITC 058 (fruit, preserved, and fruit preparations); and SITC 059 (fruit and vegetable juices, unfermented, no spirit). Also, the country recorded comparative advantages in SITC 054 in all the years, except for 2014. The product highest with the highest comparative advantage is SITC 057.

Also, SITC 057 had the most significant weighting regarding its contribution to total food exports to the EU28 as it increased from 46% in 2005 to 58% in 2017. Trailing far behind with comparative advantages are product groups SITC 034, SITC 036, and SITC 058. These product groups ranked number third, fourth, and the fifth-largest food export products to the EU28 (Table A3). In the same direction, findings by DAFF (2011); De Pablo Valenciano et al. (2017) also show that SA has been highly competitive in pear (fruits) exports to the EU28 market. They argue that the country’s competitive advantage has been driven by the TDCA signed with the EU28. Also, DAFF (2011) results reveal that SA has been competitive in products, such as fish (SITC 034) and crustaceans (SITC 036), fruits and vegetables.

The LFI findings further reveal that SA recorded high comparative disadvantages throughout the period under review in product grouping SITC 022, SITC 023, SITC 024, SITC 041, SITC 046, SITC 048, SITC 056, SITC 071, SITC 081, SITC 098 and SITC 421. This suggests that SA has not been competitive in trading with the EU28 in these product groups. The findings are partially in line with DAFF (2011) whose studies also show SA with comparative disadvantages in cereals, tobacco and sugar in trading with the EU27.

It is worth mentioning that, the country initially recorded high comparative advantage in alcoholic beverages (SITC 112), and then began to diminish, while recording comparative disadvantage between 2014 and 2015, and then 2017. Similarly, the contribution of the product grouping to exports to the EU28 also diminished from 27% in 2005 to 0.2% in 2017.

Even though the contribution of the top three (82%), top five (91%) and top ten (97%) food product groupings exported to the EU28 in 2005 merely decreased to 81%, 87%, 96% respectively in 2017, it is still substantial (Table A3). In the same direction, the HHI shows that SA’s food exports to the EU28 have been concentrated in a few products. A careful analysis of the level of value-added of the top five traded products suggests that SA widely exported fresh food, such as SITC 057, SITC 034, SITC 036, SITC 054, SITC 075 to the EU28, and the country has comparative advantages in these products. Some of these products are (tropical) commodities which the EU28 hardly produce in large quantities owing to the natural conditions of the continent as postulated by traditional trade theories.

On the contrary, SA mostly imported processed foods, such as SITC 112 (alcoholic beverages), SITC 421 (fixed vegetable fats and oils), SITC 098 (edible products and preparations), SITC 081, SITC 048 (cereal preparations, flour of fruits or vegetables), SITC 073 (chocolate, food preparations with cocoa), SITC 022 (milk and cream), and SITC 056 from the EU28, and the country has comparative disadvantages in these products (Table 2).

Additionally, it is possible that the EU’s trade policies, regarding sanitary and phytosanitary measures (SPS), non-tariff measures (NTMs) and tariff escalation, especially in semi-processed and processed foods from SA (Gebrehiwe et al., 2017) may have partially distorted trade signals and nullified the country’s efforts to boost food exports and add more value-added products with comparative advantages in trading with the EU28. Arndt and Roberts (2019) stress that there are still constrains that limits the exploitation of opportunities. Fruits and nuts are the most important leading products when SA trades with EU28 and the share of the sector on the overall agri-food export increased. However, as pointed by Adriaen et al. (2004), the EU's food market has become saturated and SA faces competition from other southern hemisphere countries with similar seasonal differences in comparison to the northern hemisphere. Also, there is a slight structural shift in food trade of SSA. The economic globalization in commodity chains contributed to the structural changes in the composition of food trade as some SSA countries moved the composition of agri-food exports from traditional to non-traditional and high-value commodities.

Agri-food trade between SA and Africa

The total value of agri-food exports from SA to African countries increased from $923 million in 2005 to its peak in 2014, with about $4.4 billion, before decreasing to $3.6 billion in 2016, and then slightly increased to about $4 Billion in 2017. Also, SA recorded a substantially positive balance of trade and TBI in trade with Africa throughout the period under review, although the TBI has decreased (Table 4).

The share of intra-African trade in total food trade merely rose from 18.4% (exports) and 16.5% (imports) in 2005 to 25.4% (exports) and 17.3% (imports) in 2017. Also, the share of intra-SADC trade in total food items rose from 19.7% (exports) and 31.5% (imports) in 2005 to 31% (exports) and 32.1% (imports) in 2017 (UNCTAD, 2019). Similarly, the share of food trade from SA to African countries rose from 23.1% (exports) and 5.7% (imports) in 2005 to 41.5%

Indicator/year		2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Export (Mill. US$)		923	867	841	1,684	1,806	3,857	4,166	4,218	4,343	4,370	3,925	3,748	3,985
Import (Mill. US$)		139	166	200	207	206	1,042	1,207	1,261	1,129	1,031	1,007	953	1,123
Balance (Mill. US$)		784	701	641	1,476	1,600	2,816	2,960	2,957	3,215	3,340	2,918	2,795	2,863
TBI		0.74	0.68	0.62	0.78	0.80	0.57	0.55	0.54	0.59	0.62	0.59	0.59	0.56
Share on export (%)*		23.14	23.21	19.96	31.01	33.04	46.86	46.21	48.24	46.74	46.18	45.92	44.48	41.46
Share on import (%)*		5.68	5.57	4.86	4.52	4.94	18.96	17.20	16.92	16.68	16.45	16.88	15.28	17.21
EX	CR3	44.9	40.1	39.7	43.7	40.9	26.2	26.0	24.9	25.9	24.7	26.1	26.1	26.7
	CR5	58.1	56.9	56.4	58.2	52.9	38.8	38.0	37.2	38.7	37.4	35.5	38.3	38.0
	CR10	76.7	76.0	75.4	77.0	71.7	61.9	60.6	61.0	62.0	59.8	57.9	59.9	60.3
	HHI	0.10	0.08	0.08	0.11	0.09	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
IM	CR3	45.1	43.8	40.2	49.0	52.9	39.9	41.8	37.9	37.6	36.9	45.5	42.0	48.1
	CR5	60.6	57.5	58.4	64.6	68.5	54.7	55.4	52.4	53.0	49.9	56.7	55.9	60.9
	CR10	82.3	78.0	81.2	82.7	87.2	79.2	79.2	74.2	76.0	73.2	78.3	77.4	81.9
	HHI	0.10	0.09	0.09	0.11	0.12	0.09	0.09	0.08	0.08	0.08	0.10	0.10	0.10
A	No. of products	36	36	38	40	40	34	35	34	34	34	34	36	37
	Export (Mill. US$)	876	820	808	1,643	1,758	3,646	3,941	3,973	4,044	3,845	3,532	3,452	3,674
	Export share (%)	94.9	94.7	96.1	97.6	97.3	94.5	94.6	94.2	93.1	88.0	90.0	92.1	92.2
	Import (Mill. US$)	33	52	70	75	60	617	740	772	614	400	336	378	425
	Import share (%)	23.4	31.1	35.1	36.0	29.1	59.2	61.3	61.2	54.4	38.8	33.4	39.7	37.9
	Balance (Mill. US$)	844	769	738	1,569	1,698	3,029	3,201	3,201	3,430	3,445	3,195	3,074	3,249
C	No. of products	3	2	1	1	0	2	1	0	1	2	1	2	0
	Export (Mill. US$)	23	16	6	6	0	34	31	0	72	347	56	28	0
	Export share (%)	2.5	1.8	0.7	0.3	0.0	0.9	0.7	0.0	1.7	7.9	1.4	0.8	0.0
	Import (Mill. US$)	18	15	6	5	0	31	27	0	59	266	40	25	0
	Import share (%)	13.3	9.1	2.9	2.6	0.0	3.0	2.3	0.0	5.2%	25.8	4.0	2.6	0.0
	Balance (Mill. US$)	4,667	751	57	63	0	3,351	3,530	0	13,139	80,332	15,853	3,662	0
D	No. of products	7	8	7	5	6	10	10	12	11	10	11	8	9
	Export (Mill. US$)	24	30	27	35	48	178	195	245	228	179	337	268	311
	Export share (%)	2.6	3.5	3.2	2.1	2.7	4.6	4.7	5.8	5.2	4.1	8.6	7.1	7.8
	Import (Mill. US$)	88	99	124	127	146	394	439	489	457	364	630	550	697
	Import share (%)	63.4	59.8	62.0	61.3	70.9	37.8	36.4	38.8	40.5	35.4	62.6	57.8	62.1
	Balance (Mill. US$)	-64	-69	-97	-92	-98	-217	-244	-244	-229	-186	-293	-283	-386

Note: * share of bilateral food trade between SA and the EU28
Source: Own composition based on UNCTAD (2019)

Table 4: Changes in food trade between SA and Africa.

(exports) and 17.2% (imports) in 2017. The SA's exports share in the region was substantially higher than the intra-Africa and intra-SADC averages, while imports were slightly below the intra-Africa and intra-SADC averages. Arguably, the increase in food trade between SA and the continent could be attributed to SA, and the African Union's efforts to stimulate local food production, value-added products, and intra-African trade. The measures may have started yielding positive results.

The 'products mapping' in Group A indicates that SA's comparative advantages in trading with African countries shifted, as it rose from 36/46 in 2005 to 40/46 in 2008, then, shrank to its lowest in 2014 with 34/46, before rising to 37/46 products in 2017. These products accounted for about 92% and 38% of the total food exports and imports respectively, between SA and all African countries (Table 4, Table A2). This indicates that the country's performance and competitiveness within the continent have been accelerated in many food products.

On the other hand, the results of the products mapping in Group D suggest that, on average, SA had comparative disadvantages in 9/46 products in trading with African countries. More detailed information on products in groups A and D are presented in Table A5. The share of these product groups also shifted during the years under study, recorded 62% and 7.8% of total imports and exports respectively in 2017 (Table 4, Table A2, Table A4). Nonetheless, the findings suggest that SA has fewer comparative disadvantages in agri-food products in trading with the African countries than the EU28 and the global markets.

A critical look at the individual product groups (Table A2) in trade with Africa shows that throughout the period under review, SA recorded comparative advantages in almost all the 46 food products (37/46) in 2017. The product groups with the highest comparative advantages in 2017 were SITC 098, SITC 112, SITC 044 (maize), SITC 022, and SITC 057 (Table A2). Interestingly, these 5 products were also among the top 5 exported products, although the share reduced from 58% in 2005 to 38% in 2017 (Table A4). Similarly, the HHI results show that food exports have been distributed more homogeneously among a broad range of products than imports (Table 3).

On the other hand, SA recorded comparative disadvantages in only a few products: SITC 001, SITC 016, SITC 034, SITC 036, SITC 061, SITC 072, SITC 057, SITC 121, and SITC 222. This suggests that SA has been competitive in trading with African countries in these product groups. A careful look at the level of value-added products of the top ten traded products shows that SA primarily exported processed food products (SITC 098, SITC 112, SITC 022, SITC 048, SITC 122, SITC 081, SITC 059), that it had comparative advantages. On the other hand, the country mainly imported fresh food products (SITC 001, SITC 034, SITC 121, SITC 057, SITC 074, SITC 011, SITC 062) from African countries.

Conclusion

This paper uses 'products mapping' tool based on TBI and Lafay index and other descriptive approaches to investigate SA's trade performance and competitiveness in agri-food with the EU28, Africa and the world. SA's agri-food trade performance products have improved since 2000. Despite SA as the EU's traditional and largest trading partner in Africa and the establishment of a preferential trade arrangement, the proportion of the EU in the total agri-food trade with the country has decreased.

The findings suggest that SA recorded more comparative advantages and leading agri-food product groups (group A) in trading with African countries (37/46) than in trading with the EU28 (13/46) in 2017. On the level of total agri-food trade, the SA reveals 19/46 leading products when trading globally. The results support the conclusion that in bilateral trade, certain products have comparative advantages in relation to African markets despite recording comparative disadvantages in relation to the EU28 market.

The number of leading product groups has remained the same in the case of trade with Africa and slightly declined in the case of trade with the EU28. The leading products mostly contribute to the positive balance of SA's agri-food trade. Contrary, the losing products (group D) are dominantly the import generating segments in the product's structure of SA's agri-food trade, negatively contributing to SA's agri-food trade balance. The groups B and C are very few in numbers and do not significantly contribute to either export or import and thus the SA's agri-food trade balance. The only exception is SA's production and trade in alcoholic beverages. This industry has shifted from A to C group when SA trades with EU28.

The comparison of SA's agri-food trade with EU28 and Africa indicated a difference in the number of leading products. Besides that, the product mapping indicates structural differences and shifts among the leading products. The product's structure within the group of leading products shows that SA has been diversified in trading with Africa while remains concentrated in export products to the EU markets. This again supports the conclusion, that emerging markets in Africa are generating opportunities for many SA agri-food industries that would otherwise not successfully compete in the EU's markets.

Recently, SA exports more processed food to Africa than it imports from the continent. The issue of reciprocal intra-regional trade and the lack of specialized trade agreements have received attention although the impact of these agreements in SA's in the region (the SADC) has not come to fruition. On the other hand, the country imports more processed products from the EU than exports, while trade with the Union has shrunk. The country should focus on producing and exporting higher value-added food products based on local raw materials.

The variability in the nature and structure of trade between the EU28 and the SA suggests that there is no apparent congruence in the growth and development of policies to improve the comparative advantages of the SA agricultural exports in the mentioned products.

Thus, policymakers in SA should continue to assess the opportunities, threats, strengths and weaknesses of the food sector to drive the effective integration of SA's agri-food industries. Also, should evaluate the existing policies regarding the exploitation of production and processing activities. For SA to realize more comparative advantages, it is imperative to improve production and trade with value-added products.

Acknowledgements

This work was supported by Internal Grant Agency of Faculty of Regional Development and International Studies, Mendel University in Brno, under no.: 2016/013 The Dynamics of EU's Economic Relations.

Corresponding authors
Ing. Ivo Zdráhal, Ph.D.
Department of Regional and Business Economics, Faculty of Regional Development and International Studies
Mendel University in Brno, Czech Republic
E-mail: ivo.zdrahal@mendelu.cz
Orcid Number 0000-0002-6631-8228

References

[1] Adriaen, D., Kleynhans T. E. and Tollens, E. (2004) "A Comparison Between South African Agricultural Exports to the European Union and those of other selected southern hemisphere countries", *Agrekon*, Vol. 43, No. 2, pp. 184-201. ISSN 2078-0400. DOI 10.1080/03031853.2004.9523644.

[2] Aiginger, K. (2000) "Specialization of European manufacturing", *Austrian Economic Quarterly*, Vol. 5, No. 2, pp. 81-92. ISSN 1098-3708.

[3] Arndt C. and Roberts, S. J. (2018) "Key issues in regional growth and integration in Southern Africa", *Development Southern Africa*, Vol. 35, No. 3, pp. 297-314, ISSN 1470-3637. DOI 10.1080/0376835X.2018.1469970.

[4] Benesova, I., Maitah, M., Smutka, L., Tomsik, K. and Ishchukova, N. (2017) "Perspectives of the Russian agricultural exports in terms of comparative advantage", *Agric. Econ. – Czech*, Vol. 63, No. 7, pp. 318-330. ISSN 1805-9295. DOI 10.17221/344/2015-AGRICECON.

[5] Benešová, I., Smutka, L. and Laputková, A. (2018) "Specifics of the Mutual Trade of the Post-Soviet Countries", *Entrepreneurial Business and Economics Review*, Vol. 7, No. 1. E-ISSN 2353-8821, ISSN 2353-883X. DOI 10.15678/EBER.2019.070102.

[6] Berends, G. (2016) "What does the EU-SADC EPA really say? An analysis of the economic partnership agreement between the European Union and Southern Africa", *South African Journal of International Affairs*, Vol. 23, No. 4, pp. 457-474. ISSN 1022-0461. DOI 10.1080/10220461.2016.1275763.

[7] Bilan Y., Mishchuk, H., Roshchyk, I. and Kmecova I. (2020) "Analysis of Intellectual Potential and its Impact on the Social and Economic Development of European Countries", *Journal of Competitiveness*, Vol.1, pp. 22-38. E-ISSN 1804-1728. ISSN 1804-171X. DOI 10.7441/joc.2020.01.02

[8] Blažková, I. and Chmelíková, G. (2016) "The Influence of Market Concentration on the Development of Newly Born Business in the Czech Republic", *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, Vol. 64, No. 3, pp. 929-938. ISSN 1211-8516. DOI 10.11118/actaun201664030929.

[9] Blížkovský, P., Grega, L. and Verter, N. (2018) "Towards a common agricultural policy in Africa?", *Agric. Econ. – Czech*, Vol. 64, No. 7, pp. 301-315. ISSN 0139-570X. DOI 10.17221/310/2016-AGRICECON.

[10] Bojnec, Š. and Fertő, I. (2015) "Agri-food export competitiveness in European Union countries", *JCMS: Journal of Common Market Studies*, Vol. 53, No. 3, pp. 476-492. E-ISSN 1468-5965. DOI *10.1111/jcms.12215*.

[11] Brodzicki, T. and Kwiatkowski, J. (2018) "An Empirical Investigation into the Role of Technology Gap in the Trade Relations of the EU Member States", *Entrepreneurial Business and Economics Review*, Vol. 6, No. 2, pp. 111-135. E-ISSN 2353-8821, ISSN 2353-883X. DOI 10.15678/EBER.2018.060206

[12] Cieślik, A., Qu, Y. and Qu, T. (2018) "Innovations and Export Performance: Firm Level Evidence from China", E*ntrepreneurial Business and Economics Review*, Vol. 6, No. 4, pp. 27-47. E-ISSN 2353-8821, ISSN 2353-883X. DOI 10.15678/EBER.2018.060402

[13] Department of Agriculture, Forestry and Fisheries of South Africa (DAFF) (2011) "*Competitiveness of selected South African agricultural product exports in the EU-27"*, Economic Services and Directorate, International Trade, South Africa. [Online]. Available: https://1url.cz/6MgC8 [Accessed: 14 May 2019].

[14] De Pablo Valenciano, J., Uriarte, M. M. C. and Battistuzzi, M. A. G. (2017) "South Africa's competitiveness against its main competitors in the market of pears imported by the EU28", *Agrekon*, Vol. 56, No. 2, pp. 191-204. ISSN 2078-0400. DOI 10.1080/03031853.2017.1297723.

[15] DTI – Department of Trade and Industry Republic of South Africa (2020) "*African Regional Integration*". [Online]. Available: https://bit.ly/33fYsYX [Accessed: 15 Feb 2020].

[16] Economist Intelligence Unit (EIU) (2019) *"Global food security index 2019: Exploring challenges & developing solutions"*, EIU. [Online]. Available at: https://bit.ly/2THEbs8 [Accessed: 11 March 2020].

[17] Esquivias, M. A. (2017) "The Change of Comparative Advantage of Agricultural Activities in East Java Within the Context of Asean Economic Integration", *AGRIS on-line Papers in Economics and Informatics*, Vol. 9, No. 1, pp. 33-47. ISSN 1804-1930. DOI 10.7160/aol.2017.090103.

[18] European Commission (2016) "Economic Partnership Agreement between the European Union and its Member States, of the one part, and the SADC EPA States, of the other part", *Official Journal of the European Union*, L 250/2108.

[19] European Commission (2019) "*Agri-food trade factsheet: European Union and South Africa*". [Online]. Available: https://bit.ly/2McDx4j [Accessed: 27 May 2019].

[20] Fertő, I. (2008) "The evolution of agri-food trade patterns in Central European countries", *Post-Communist Economies*, Vol. 20, No. 1, pp. 1-10. ISSN 1463-1377. DOI 10.1080/14631370701865680.

[21] Food and Agriculture Organization of the United Nations (FAO) (2011) "*Why has Africa become a net food importer? Explaining Africa agricultural and food trade deficits*". [Online]. Available: https://bit.ly/2IzRtAy [Accessed: 20 Jan 2019].

[22] Food and Agriculture Organization of the United Nations (FAO) (2019) "F*AOSTAT database*", [Online]. Available: http://bit.ly/NmQzZf [Accessed: 17 June 2019].

[23] Food Security Information Network (FSIN) (2019) "*Global report on food crises 2019*", FSIN. [Online]. Available: http://fsinplatform.org/ [Accessed: 27 July 2019].

[24] Fujita, M., Krugman, P. and Venables, A. (1999) "*The Spatial Economy: Cities, Regions, and International Trade*", Cambridge, MA: MIT Press, 382 p. ISBN 9780262561471. DOI 10.7551/mitpress/6389.001.0001.

[25] Gebrehiwet, Y., Ngqangweni, S. and Kirsten, J. F. (2007) “Quantifying the trade effect of sanitary and phytosanitary regulations of OECD countries on South African food exports”, *Agrekon*, Vol. 46, No. 1, pp. 1-17. ISSN 2078-0400. DOI 10.1080/03031853.2007.9523759

[26] Ishchukova, N. and Smutka, L. (2014) “Revealed" comparative advantage: products mapping of the Russian agricultural exports in relation to individual regions“, *Acta Scientiarum Polonorum. Oeconomia*, Vol. 13, No. 1, pp. 45-61. ISSN 1644-0757.

[27] Kostiukevych, R., Mishchuk, H., Zhidebekkyzy, A., Nakonieczny, J. and Akimov, O. (2020) "The impact of European integration processes on the investment potential and institutional maturity of rural communities", *Economics and Sociology*, Vol. 13, No. 3, pp. 46-63. E-ISSN 2306-3459, ISSN 2071-789X. DOI 10.14254/2071-789X.2020/13-3/3.

[28] Krugman, P. (1987) “The narrow mowing band, the Dutch disease, and the competitive consequences of Mrs. Thatcher: notes on trade in the presence of dynamic scale economies”, *Journal of Development Economics*, Vol. 27, No. 1, pp. 41-54. ISSN 0304-3878. DOI 10.1016/0304-3878(87)90005-8.

[29] Krugman, P. (1991) “*Geography and Trade*”, Cambridge, MA: MIT Press, p. 156. ISBN 9780262610865.

[30] Lafay, G. (1992) “The measurement of revealed comparative advantages”, In: DAGENAIS, M. G. & MUET, P.-A. (eds.), "*International Trade Modelling"*. London: Chapman & Hall. ISBN 978-0412450006.

[31] Larsén, M. F. (2007) “Trade Negotiations between the EU and South Africa: A Three-Level Game”, *JCMS: Journal of Common Market Studies*, Vol. 45, No. 4, pp. 857-881. ISSN 1468-5965. DOI 10.1111/j.1468-5965.2007.00751.x.

[32] Lucas, R. (1988) “On the mechanics of economic development”, *Journal of Monetary Economics*, Vol. 22, No. 1, pp. 3-22. ISSN 0304-3932. DOI 10.1016/0304-3932(88)90168-7.

[33] Maris, M. (2019) "Structural and productivity shift of industries in Slovakia and Czech Republic: A comparative study", *Journal of International Studies*, Vol. 12, No. 1, pp. 313-323. E-ISSN 2306-3483, ISSN 2071-8330. DOI 10.14254/2071-8330.2019/12-1/21.

[34] Nazarczuk, J., Umiński, S. and Gawlikowska-Hueckel, K. (2018) "The Role of Specialisation in the Export Success of Polish Counties in 2004-2015", *Entrepreneurial Business and Economics Review*, Vol. 6 No. 2, pp. 91-109. E-ISSN 2353-8821, ISSN 2353-883X. DOI 10.15678/EBER.2018.060205

[35] Ortikov, A., Smutka, L. and Benešová, I. (2019) “Competitiveness of Uzbek agrarian foreign trade – different regional trade blocs and the most significant trade partners”, *Journal of International Studies*, Vol. 12, No. 4, pp. 177-194. ISSN 2306-3483. DOI 10.14254/2071-8330.2019/12-4/12.

[36] Porter, M. E. (1990) “Competitive advantage of nations”, *Harvard Business Review,* Vol. 68, No. 2, pp. 73-93. ISSN 0017-8012.

[37] Reis, J. G. and Farole, T. (2012) “*Trade Competitiveness Diagnostic Toolkit*”, Washington, DC, World Bank Group. DOI 10.1596/978-0-8213-8937-9.

[38] Sandrey, R., Punt, C., Jensen, H. G. and Vink, N. (2011) "Agricultural Trade and Employment in South Africa“, *OECD Trade Policy Papers*, No. 130, OECD Publishing, Paris. [Online]. Available: https://bit.ly/2TExdEa [Accessed: 01 Dec 2019]. E-ISSN 18166873. DOI 10.1787/5kg3nh58nvq1-en.

[39] Sanidas, E. and Shin, Y. (2010) “*Comparison of Revealed Comparative Advantage Indices with Application to Trade Tendencies of East Asian Countries*”, Seoul National University. [Online]. Available: https://bit.ly/39LOzVe [Accessed: 20 April 2019].

[40] Smutka, L., Rovný, P., Pulkrábek, J. and Horská, E. (2016) “Foreign Trade in Sugar (HS 1701) and Sugar Containing Products (HS17) – Comparison of the Czech and Slovak Republic During the Last Two Decades”, *Listy Cukrovarnicke a Reparske*, Vol. 132, No. 7-8, pp. 234-240. ISSN 1805-9708.

[41] Smutka, L., Maitah, M. and Svatoš, M. (2018) "Changes in the Czech agrarian foreign trade competitiveness – different groups of partners' specifics", *Agric. Econ. – Czech*, Vol. 64, No. 9, pp. 399-411. ISSN 1805-9295. DOI 10.17221/399/2016-AGRICECON.

[42] Smutka, L., Maitah, M. and Svatoš, M. (2019) "The Czech Agrarian Trade Comparative Advantages Distribution Based on Value and Volume Approach", *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, Vol. 67, No. 6, pp. 1613-1625. ISSN 2464-8310. DOI 10.11118/actaun201967061613.

[43] UNCTAD (2019) "*World statistical database*" [Online]. Available: https://bit.ly/21GbfKX. [Accessed: 25 Feb 2019].

[44] Vink, N. and van Rooyen, J. (2009) "Perspectives on the Performance of Agriculture in South Africa since 1994 and implications for its Role in Achieving Sustainable Food Security", Working Papers 108005, University of Pretoria, Department of Agricultural Economics, Extension and Rural Development. [Online]. Available: https://bit.ly/2IFAo8g [Accessed: 10 Dec 2019]. DOI 10.22004/ag.econ.108005

[45] Verter, N., Zdráhal, I., Bečvářová, V. and Grega, L. (2020) "'Products mapping' and trade in agri-food products between Nigeria and the EU28", *Agric. Econ. – Czech*, Vol. 66, No. 1, pp. 34-45. ISSN 0139-570X. DOI 10.17221/145/2019-AGRICECON

[46] Widodo, T. (2008) "Product mapping and dynamic shift in the patterns of comparative advantage: Could India Catch Up China?", *Journal of Chinese Economic and Foreign Trade Studies*, Vol. 1, No. 3, pp. 200-213. ISSN 1754-4408. DOI 10.1108/17544400810912365.

[47] Widodo, T. (2009) "Comparative Advantage: Theory, Empirical Measures and Case Studies", *Review of Economic and Business Studies*, Vol. 2, No. 2, pp. 57-81. ISSN 1843-763X.

[48] World Bank (2019) "*World development indicators*", [Online]. Available at: http://bit.ly/1Cd8EkQ. [Accessed: 28 June 2019].

[49] Zaghini, A. (2003) "Trade advantages and specialization dynamics in acceding countries", Working paper No. 249, Frankfurt am Main, Germany: European Central Bank, pp. 4-15.

[50] Zdráhal, I., Verter, N., Daňková, B. and Kuchtík, J. (2019) "'Products mapping' and agri-food trade between Nigeria and ECOWAS member countries", *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, Vol. 67, No. 5, pp. 1379-1392. ISSN 1211-8516. DOI 10.11118/actaun201967051379.

[51] Yang, J., Černevičiūtė, J., Strazdas, R. and Tvaronavičienė, M. (2019) "Stimulation of cultural and creative industries clusters development: A case study from China", *Terra Economicus*, Vol. 17, No. 3, pp. 169-184. ISSN 24104531.

Appendix

SITC code	Food product	SITC code	Food product
001	Live animals	057	Fruit, nuts excl. oil nuts
011	Bovine Meat	058	Fruit, preserved, prepared
012	Other meat, other offal	059	Fruit, vegetable juices
016	Meat, ed. offl., dry, slt, smk	061	Sugars, molasses, honey
017	Meat, offl. Prdd, nes	062	Sugar, confectionery
022	Milk and cream	071	Coffee, coffee substitutes
023	Butter, other fat of milk	072	Cocoa
024	Cheese and curd	073	Chocolate, oth. cocoa prep.
025	Eggs, birds, yolks, albumin	074	Tea and mate
034	Fish, fresh, chilled, frozn	075	Spices
035	Fish, dried, salted, smoked	081	Animal feed stuff
036	Crustaceans, Molluscs	091	Margarine and shorten
037	Fish etc. prepd, prsvd. nes	098	Edible prod. prepetns, nes
041	Wheat, Meslin, Unmilled	111	Non-alcohol. beverage
042	Rice	112	Alcoholic Beverages
043	Barley, unmilled	121	Tobacco, unmanufactured
044	Maize unmilled	122	Tobacco, manufactured
045	Other cereals, unmilled	222	Oil seeds and oleaginous fruits (excl. flour)
046	Meal, Flour of wheat, msln	223	Oil seeds, oleaginous fruits (incl. flour, n.e.s.)
047	Other cereal meal, flours	411	Animal oils and fats
048	Cereal preparations	421	Fixed veg. fat, oils, soft
054	Vegetables	422	Fixed veg. fat, oils, other
056	Vegetables, prpd, prsvd, nes	431	Animal, veg. Fats, oils, nes.

Source: SITC rev.3

Table A1: Sectors and their numeric designations (SITC rev.3, 3-digit code).

SITC	2005	2010	2015	2017	2005	2010	2015	2017	2005	2010	2015	2017
South Africa's Agri-food trade with												
World					**EU28**				**Africa**			
001	-0.15	-1.20	-1.38	-2.10	-0.16	-0.10	-0.08	-0.07	0.11	-0.56	-0.70	-1.17
011	-0.74	-0.71	0.74	0.69	0.00	0.00	0.00	0.00	0.05	-0.35	-0.01	0.01
012	-6.83	-3.74	-5.54	-6.51	-0.28	-0.27	-2.21	-1.19	0.07	0.43	0.53	0.54
016	-0.02	-0.05	-0.05	-0.04	-0.01	-0.01	-0.01	-0.02	0.01	-0.02	-0.02	-0.01
017	0.12	0.19	0.45	0.40	-0.03	-0.01	0.02	0.00	-0.02	0.14	0.23	0.21
022	-0.96	0.94	0.71	0.67	-0.39	-0.21	-0.50	-0.47	0.24	1.00	1.02	0.99
023	-0.16	0.00	-0.11	-0.26	-0.12	-0.04	-0.05	-0.12	0.02	0.08	0.06	0.05
024	-0.45	-0.21	-0.43	-0.37	-0.18	-0.17	-0.35	-0.34	0.02	0.15	0.18	0.19
025	-0.02	0.25	0.33	0.16	0.00	-0.01	-0.02	-0.02	-0.01	0.16	0.21	0.12
034	3.89	0.60	0.18	-0.48	1.94	0.92	0.80	0.76	0.07	-0.51	-0.52	-0.65
035	0.05	0.06	0.00	0.03	0.00	0.00	0.00	0.00	0.04	0.00	-0.02	0.00
036	1.15	1.08	0.72	0.23	0.74	0.39	0.28	0.42	-0.14	-0.05	-0.02	-0.04
037	-0.91	-1.76	-1.88	-0.69	0.03	0.07	0.05	0.08	0.04	-0.07	-0.14	0.03
041	-5.52	-3.54	-5.64	-3.81	-0.50	-1.06	-1.00	-1.22	0.07	0.31	0.51	0.15
042	-7.00	-5.38	-5.33	-5.81	-0.03	0.00	0.00	-0.01	0.06	0.40	0.39	0.30
043	-0.58	-0.22	-0.30	-0.14	-0.08	-0.02	0.00	0.00	0.00	0.00	0.00	0.00
044	4.69	4.01	-0.09	2.31	0.01	0.11	0.02	-0.02	2.62	1.42	0.96	1.37
045	-0.26	0.04	-0.23	-0.14	0.00	-0.02	-0.09	0.00	0.02	0.09	0.06	0.04
046	0.04	0.42	0.24	0.01	-0.04	-0.02	-0.04	-0.02	0.07	0.28	0.19	0.03
047	2.02	0.66	0.96	0.89	-0.01	0.00	0.00	0.00	1.24	0.39	0.57	0.53
048	-1.64	-0.18	-0.18	-0.24	-0.49	-0.57	-0.69	-0.73	0.17	0.89	0.94	0.90
054	0.06	0.30	0.83	1.40	0.29	0.09	0.04	0.13	0.22	0.76	0.73	0.74
056	-0.79	-0.53	-0.22	-0.33	-0.21	-0.36	-0.20	-0.28	0.03	0.34	0.37	0.36
057	22.13	20.14	25.97	26.52	9.31	6.16	7.24	7.25	0.35	1.00	1.16	0.79
058	2.96	1.84	1.45	1.25	0.90	0.48	0.34	0.31	0.04	0.11	0.03	0.04
059	1.75	2.00	1.73	1.15	0.37	0.20	0.26	0.11	0.27	0.90	0.95	0.76
061	4.51	-0.54	-3.20	-4.44	-0.10	-0.10	-0.10	0.15	0.93	0.25	-1.01	-0.84
062	-1.25	-0.44	-0.02	-0.02	-0.09	-0.04	-0.06	-0.05	0.04	0.02	0.17	0.15
071	-1.38	-1.21	-2.02	-1.79	-0.24	-0.22	-0.51	-0.46	-0.03	0.11	0.03	0.05
072	-1.07	-1.02	-0.88	-0.63	-0.07	-0.04	-0.04	-0.05	-0.29	-0.23	-0.15	-0.09
073	-0.51	0.03	-0.82	-0.65	-0.40	-0.27	-0.60	-0.60	0.18	0.15	0.24	0.23
074	-0.64	-0.49	-0.49	-0.47	0.01	-0.02	-0.05	-0.03	-0.35	-0.22	-0.14	-0.16
075	-0.35	-0.22	-0.39	-0.25	0.03	0.00	0.00	0.00	-0.03	0.21	0.25	0.24
081	-5.11	-5.80	-2.85	-2.84	-0.46	-0.65	-0.73	-0.73	-0.03	0.21	0.61	0.34
091	-0.33	0.06	0.13	0.17	-0.12	-0.02	-0.04	-0.03	0.08	0.24	0.22	0.20
098	-2.42	0.60	0.59	0.90	-1.18	-0.88	-1.08	-1.02	0.71	1.94	2.14	2.08
111	-0.17	0.13	0.45	0.32	-0.41	-0.29	-0.36	-0.30	0.33	0.43	0.79	0.68
112	6.22	3.71	4.21	3.48	2.85	0.65	-0.01	-0.07	0.70	0.80	1.55	1.33
121	-1.71	-2.34	-1.75	-1.52	-0.15	-0.05	-0.05	-0.02	-0.42	-0.39	-0.44	-0.55
122	1.76	2.45	1.50	0.60	-0.22	-0.17	-0.06	-0.06	0.60	1.30	1.00	0.82
222	-0.15	0.37	-1.14	-0.75	0.09	-0.01	-0.09	-0.15	-0.18	-0.08	-0.04	-0.08
223	0.11	0.24	0.09	0.10	0.00	0.00	0.00	0.00	0.09	0.15	0.08	0.09
411	-0.45	-0.25	0.05	-0.03	-0.01	-0.01	-0.01	0.01	0.02	-0.01	0.01	0.02
421	-4.40	-4.13	-1.94	-2.74	-0.36	-1.94	-1.37	-1.60	0.18	1.14	0.93	0.72
422	-4.17	-5.04	-3.88	-4.56	-0.07	-0.02	-0.03	-0.02	0.01	0.03	0.07	0.09
431	-1.31	-1.14	-0.55	0.29	-0.07	-0.03	0.01	0.10	0.03	0.05	0.11	0.09
Total	**0.00**	**0.00**	**0.00**	**0.00**	**10.08**	**1.42**	**-1.38**	**-0.37**	**8.23**	**13.39**	**14.06**	**11.68**

Source: Own composition based on UNCTAD (2019)

Table A2: South Africa's comparative advantage (LFI) in agri-food products with World, EU28 and Africa.

Indicator/year	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Exports to the EU28 (%)													
[057] Fruits and nuts	45.5	41.8	45.3	42.3	45.6	48.4	49.2	52.7	53.4	54.0	58.5	57.9	58.1
[112] Alcoholic beverages	27.0	26.4	26.1	25.3	27.5	25.7	23.8	23.3	24.8	21.0	18.5	16.8	16.3
[034] Fish, fresh (live or dead)	9.5	9.0	8.7	10.2	7.8	7.3	7.9	7.4	6.3	7.6	6.7	7.1	6.5
[036] Crustaceans, mollusks, etc	3.9	4.5	4.4	3.8	3.0	3.3	3.2	2.6	1.6	2.0	2.4	3.2	3.4
[058] Fruit, preserved, and fruit	5.1	5.3	4.3	4.1	4.5	4.3	3.9	4.0	3.4	3.2	3.2	3.0	2.9
[061] Sugar, molasses/oney	0.3	0.1	0.9	0.1	0.1	0.0	0.1	0.0	0.1	0.2	0.1	0.1	2.4
[054] Vegetables	1.8	1.8	1.6	1.9	1.3	1.2	1.5	1.5	1.5	1.4	1.5	2.2	2.3
[059] Fruit and vegetable juices	1.9	3.4	1.5	1.5	1.6	2.0	2.0	1.9	2.4	2.4	2.5	2.1	1.6
[431] Animal or veg. oils & fats	0.2	0.2	0.2	0.2	0.2	0.1	0.3	0.3	0.3	0.4	0.4	0.4	1.1
[056] Vegetables, roots, tubers, etc	0.6	0.6	0.6	0.8	0.9	0.9	1.1	0.9	0.9	1.1	1.0	0.9	1.0
Total top 10 products	95.5	93.1	93.8	90.3	92.4	93.0	93.0	94.5	94.6	93.2	94.8	93.5	95.5
Imports from the EU28 (%)													
[112] Alcoholic beverages	25.8	27.1	34.7	32.9	33.0	23.2	20.4	19.2	22.1	17.1	16.7	15.4	16.3
[421] Fixed vegetable fats & oils,	3.5	2.6	3.5	4.4	2.5	16.8	19.7	16.4	14.3	13.2	9.9	13.0	12.2
[012] Other meat and edible meat	3.8	5.2	4.1	3.3	2.9	4.9	10.1	13.8	14.0	14.7	16.3	17.9	9.4
[041] Wheat (incl. spelt) & meslin	4.6	7.8	1.3	7.1	15.2	9.2	4.3	1.5	1.8	5.5	7.2	5.2	9.2
[098] Edible products and prep.	12.6	11.9	11.5	11.2	10.6	8.5	8.4	8.5	9.3	8.4	8.7	8.6	8.4
[081] Feeding stuff for animals	4.6	4.7	5.1	5.8	5.5	5.7	5.7	5.4	5.4	5.3	6.0	5.1	6.1
[048] Cereal preparations	4.8	5.2	5.6	5.5	4.5	5.0	5.7	5.1	5.1	5.5	5.0	5.6	5.7
[073] Chocolate, food prep.	4.1	4.5	2.7	2.4	2.2	2.4	2.8	3.6	4.1	3.9	4.3	3.7	4.6
[071] Coffee and coffee	2.4	2.8	2.0	2.4	1.9	2.0	2.2	3.0	3.4	3.3	3.7	3.3	3.6
[022] Milk and cream	3.7	2.3	3.2	2.8	1.7	1.8	2.5	3.1	2.7	3.2	3.6	2.5	3.5
Total top 10 products	69.8	74.0	73.7	77.7	79.8	79.6	81.8	79.7	82.2	79.9	81.2	80.2	79.0

Source: Own composition based on UNCTAD (2019)

Table A3: Share of top 10 food products (% of total food trade) between SA and the EU28.

Indicator/year	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Exports within Africa (%)													
[098] Edible products and prep.	6.8	10.3	12.3	7.5	7.9	8.9	9.4	9.1	8.9	9.4	10.0	9.8	10.8
[112] Alcoholic beverages	6.4	6.5	9.2	7.0	7.0	8.3	9.3	8.6	8.6	8.8	9.1	7.9	8.9
[044] Maize	24.0	13.9	2.3	27.7	23.5	6.4	4.4	4.6	6.5	6.3	4.5	8.4	6.9
[057] Fruits and nuts	4.4	6.5	9.1	5.4	4.9	5.3	5.2	5.7	6.4	6.4	7.1	6.5	6.4
[022] Milk and cream	2.2	2.9	3.3	2.5	3.1	4.5	4.5	4.9	5.2	5.1	4.6	4.1	5.0
[048] Cereal preparations	1.8	2.3	2.9	2.3	2.8	4.1	4.5	4.5	4.5	4.5	4.4	4.5	4.8
[061] Sugar, molasses/honey	9.6	15.9	18.2	8.5	9.5	9.1	7.4	7.2	8.5	6.5	4.4	3.9	4.7
[054] Vegetables	2.5	3.2	3.4	2.1	1.8	4.0	4.0	3.9	3.8	3.9	3.8	4.4	4.4
[122] Tobacco, manufactured	6.1	10.3	7.6	3.3	4.7	6.2	5.5	5.7	4.9	4.2	4.7	4.1	4.3
[081] Feeding stuff for animals	1.4	1.4	1.8	1.5	1.3	3.2	3.2	3.7	4.1	3.9	4.8	5.7	4.2
Total top 10 products	65.2	73.1	70.1	67.8	66.5	59.8	57.3	57.9	61.2	58.9	57.3	59.3	60.3
Imports within Africa (%)													
[061] Sugar, molasses/honey	4.2	3.0	4.6	3.0	4.0	19.7	19.3	17.6	17.2	20.1	24.4	24.0	21.4
[001] Live animals	0.6	0.6	0.3	0.4	0.2	8.4	9.6	6.5	9.0	5.4	11.7	7.4	16.6
[034] Fish, fresh (live or dead)	1.3	1.8	1.6	1.4	2.3	8.6	7.2	8.0	8.8	8.8	9.4	10.6	10.1
[121] Tobacco, unmanufactured	19.2	20.3	18.1	23.8	23.2	4.8	5.5	3.8	3.2	6.6	5.5	7.5	6.8
[057] Fruits and nuts	4.8	7.0	6.8	6.6	5.9	2.1	2.2	3.1	3.9	3.8	5.0	5.5	5.9
[074] Tea and mate	14.9	15.4	10.5	11.4	16.9	3.9	3.0	3.2	3.9	3.9	3.5	4.1	3.7
[081] Feeding stuff for animals	6.8	6.7	11.6	13.8	12.8	5.5	5.6	7.3	6.3	6.4	5.5	6.5	6.0
[011] Meat of bovine animals	0.0	0.1	0.0	0.1	0.1	6.5	6.4	7.1	6.5	5.7	4.0	3.8	3.3
[062] Sugar confectionery	1.1	2.9	3.1	1.9	1.4	4.6	4.1	3.7	5.2	4.3	3.0	2.8	2.8
Total top 10 products	52.9	58.7	57.0	63.3	67.0	75.6	75.8	72.6	75.4	73.0	77.8	77.4	81.9
Total top 10 products	69.8	74.0	73.7	77.7	79.8	79.6	81.8	79.7	82.2	79.9	81.2	80.2	79.0

Source: Own composition based on UNCTAD (2019)

Table A4: Share of top 10 food products (% of total food trade) between SA and Africa.

Group A (leading products)					Group D (losing products)				
SITC	Export	SonEX	Import	SonIM	SITC	Export	SonEX	Import	SonIM
	1000 USD	%	1000 USD	%		1000 USD	%	1000 USD	%
Trade in Food Products between South Africa and the EU28									
2005									
[057]	800,677	45.51	7,636	1.37	[098]	13,745	0.78	69,845	12.55
[112]	474,160	26.95	143,305	25.75	[048]	1,410	0.08	26,522	4.77
[034]	167,428	9.52	1,774	0.32	[041]	0	0.00	25,765	4.63
[058]	89,291	5.08	7,977	1.43	[081]	1,971	0.11	25,327	4.55
[036]	67,650	3.85	3,091	0.56	[111]	4,344	0.25	24,010	4.31
others	88,183	5.01	10,943	1.97	others	50,256	2.86	210,117	37.76
Sum	*1,687,389*	*95.91*	*174,725*	*31.40*	*Sum*	*71,725*	*4.08*	*381,585*	*68.57*
2017									
[057]	1,484,329	58.10	25,241	1.40	[421]	727	0.00	217,140	12.20
[034]	164,890	6.50	9,563	0.50	[012]	8,434	0.30	167,094	9.40
[036]	87,914	3.40	2,543	0.10	[041]	0	0.00	164,856	9.20
[058]	74,758	2.90	8,995	0.50	[098]	18,709	0.70	150,576	8.40
[061]	60,423	2.40	21,147	1.20	[081]	14,742	0.60	108,495	6.10
others	1,872,315	73.26	67,490	3.78	others	54,917	2.15	561,519	31.45
Sum	*2,035,060*	*79.63*	*118,777*	*6.65*	*Sum*	*97,530*	*3.82*	*1,369,679*	*76.71*
Trade in Food Products between South Africa and Africa									
2005									
[044]	221,604	24.00	92	0.07	[121]	8,433	0.91	26,781	19.24
[047]	104,832	11.35	8	0.01	[074]	4,042	0.44	20,742	14.90
[061]	88,185	9.55	5,796	4.16	[072]	752	0.08	15,321	11.00
[098]	62,738	6.79	1,850	1.33	[222]	4,759	0.52	12,119	8.70
[112]	59,166	6.41	196	0.14	[036]	2,239	0.24	8,772	6.30
others	339,892	36.81	24,591	17.66	others	3,561	0.39	4,497	3.23
Sum	*876,418*	*94.92*	*32,533*	*23.37*	*Sum*	*23,786*	*2.58*	*88,233*	*63.37*
2017									
[098]	432,156	10.84	11,616	1.03	[061]	185,665	4.70	240,348	21.40
[112]	355,749	8.93	61,097	5.44	[001]	41,399	1.00	186,006	16.60
[044]	275,213	6.91	1,844	0.16	[034]	37,325	0.90	113,736	10.10
[057]	253,445	6.36	65,692	5.85	[121]	2,227	0.10	75,905	6.80
[022]	198,215	4.97	853	0.08	[074]	28,952	0.70	40,978	3.70
others	2,159,229	54.18	284,070	25.30	others	15,577	0.39	40,507	3.61
Sum	*3,674,006*	*92.19*	*425,172*	*37.87*	*Sum*	*311,145*	*7.81*	*697,480*	*62.13*

Notes: SonEX – share on exports, SonIM – share on imports
Source: Own composition based on UNCTAD (2019)

Table A5: Results of Product Mapping for South Africa's agri-food trade with EU28 and Africa (2005, 2017).

Agris on-line Papers in Economics and Informatics

Volume XII Number 4, 2020

Agris on-line Papers in Economics and Informatics

The international reviewed scientific journal issued by the Faculty of Economics and Management of the Czech University of Life Sciences Prague.

http://online.agris.cz

ISSN 1804-1930